

Opp av lenestolen?

Psykologisk realisme og vitenskapens relevans for etikken

Hanne Rennesund Tallaksen



Masteroppgave i filosofi

Veileder: Bjørn Ramberg

Institutt for filosofi, ide- og kunsthistorie og klassiske språk

UNIVERSITETET I OSLO

Vår 2011

Sammendrag

Hvordan lar filosofiens moralteorier seg påvirke av empirisk basert kunnskap om hvem vi er? Denne oppgaven undersøker én type moralsk naturalisme, nemlig naturalisme som kompatibilisme. Dette innebærer å konstruere teorier i samsvar med hva empiriske vitenskaper har å fortelle om hva et menneske er. En kompatibilists moralpsykologi vil derfor være psykologisk realistisk, det vil si, innenfor de rammene for hva som er realiserbart for vesener som oss. Å ta høyde for empiriske data i moralfilosofien synes å ha betydning både for metode, og for hvilke elementer som anses som sentrale i en moralteori. Dette belyses ved å se på en konkret debatt: debatten om dydenes realiserbarhet og empiriske holdbarhet. Med utgangspunkt i dydsetikken vil vi også se at det finnes grenser for i hvor stor grad man kan appellere til naturen i moralteori. Målet for oppgaven er altså å undersøke hva en mer empirisk sensitiv moralfilosofi vil medføre av forpliktelser til metode og i teori.

Innhold

Innledning.....	7
1. Naturalisme og kompatibilisme.....	9
1.1 Prinz om empirisk filosofi	18
1.2 Flanagan om minimal psykologisk realisme.....	22
2. Psykologisk realisme i praksis	31
2.1 Dydsetikk og empiri.....	34
2.2 Dydene som karaktertrekk	37
2.3 John Doris' kritikk av dydsetikken	40
2.4 Annas' respons.....	44
2.5 Arpalys respons.....	52
2.6 Solomons respons	55
2.7 Noen lærdommer	60
3. Gjør dydene oss lykkelige? Om naturalismens grenser	65
Konklusjon	72
Kildeliste	74

Innledning

I sin selvbiografi beskriver den amerikanske filosofen Stanley Cavell hvordan han alltid lar det bli igjen en liten slump mat på tallerkenen etter et måltid. Dette gjør han ikke fordi han ikke orker mer, eller av hensyn til etikette. Han gjør det for å minne seg selv om at han er fri: "I recognize that to this day I unfailingly at the end of a meal leave some portion of food, if sometimes quite small, on my dish – as if to reassure myself that I am free (Cavell 2010: 107). Det er en liten handling, med stor betydning. Ett enkelt rituale for å bevise at mennesket *ikke bare* er natur. Ofte kommer denne erkjennelsen av oss selv som noe over og vedsiden av naturen til syne, ikke bare gjennom oppfattelsen av oss selv som frie, men gjennom hvordan vi bringer noe normativt inn i verden. Gitt alle våre drifter og fysiologiske behov synes det som om moralen befinner seg i en helt annen sfære. Det er noe besnærende ved tanken på moralen som noe "over naturen". Kanskje trenger vi å minnes på at mennesket *også bare* er et dyr. En måte å gjøre dette på er å se på moralen som sensitiv for hvordan vi er konstruert som naturvesener.

Dette er en oppgave som utforsker én side av moralsk naturalisme. Mer presist er jeg interessert i å betrakte og argumentere for moralsk kompatibilisme. Kompatibilisme innenfor rammene av denne oppgaven innebærer én bestemt naturalistisk forpliktelse, nemlig å konstruere (moral)teorier som er i samsvar med hva vitenskapen kan fortelle om hvem vi er. Det vil si, om hvilke psykologiske strukturer, mekanismer og elementer som sammen utgjør et menneskesinn. Naturalisme som reduksjonisme er derfor ikke et tema her. Jeg vil ikke ta stilling til hvorvidt moralske egenskaper lar seg identifisere med naturlige egenskaper. Jeg forutsetter riktignok at moralske fakta er fakta av en type som er sensitive for empiriske hensyn, men jeg nøyer meg med å skille dette fra de som mener at moralen er noe *sui generis* og ureduserbart (Miller 2003: 4). Jeg har valgt å fokusere på samspillet mellom moral og én gren av vitenskapen, nemlig kognitive vitenskaper. Naturalisme i denne oppgaven er altså å forstå som kompatibilisme, og videre vil jeg bringe denne kompatibilismen inn i et rammeverk av kunnskap om hvordan vi er psykologisk oppbygd og utstyrt. Naturalisme i denne oppgaven kommer altså til uttrykk gjennom det som kan kalles psykologisk realisme: En forpliktelse til å konstruere teorier og bedrive moralfilosofi innenfor de mulighetskriteriene som gis oss av kognitive vitenskaper. Jeg tror at et tettere

forhold mellom moral og empiri vil ha konsekvenser for filosofi på to nivåer: både metodisk, og for innholdet i adekvate moralteorier.

For å nærme oss noen konklusjoner om hva en empirisk sensitiv moralfilosofi og moralteori vil innebære tar jeg tak i en konkret debatt. Dydsetikken har de siste tjue årene vært gjenstand for en empirisk basert kritikk, nemlig anklager om at dydene ikke er psykologisk realistiske (det vil si, ikke holdbare i lys av empiriske studier). Debatten om dydsetikken belyser både metodologiske utfordringer, og hvilke hensyn vi må ta for å kunne bygge teorier som er psykologisk realistiske. Dydsetikken gir seg ut for å være naturalistisk, og i den sammenhengen tar jeg opp en annen empirisk basert kritikk. Kan dydsetikerne egentlig finne fundament for dydene i betraktninger om hvordan vi er fra naturens side? Denne kritikken peker hen mot naturalismens grenser. Kunnskap om hvem vi er hjelper oss bare et stykke på veien mot adekvate moralteorier. Hvordan vi bør leve og hvordan vi organiserer samfunnet på best mulig måte er temaer som krever selvstendig redegjøring utover betraktninger om vår naturlige beskaffenhet. Jeg håper, og tror, at gjennom å studere en konkret debatt vil man klarere kunne se hvilke utfordringer som ligger i å bedrive, det som senere i oppgaven vil bli presentert som, empirisk filosofi. Dessuten tror jeg at vi ikke bare vil oppdage utfordringer, men også se muligheter for å skape spennende og interessante moralteorier.

Til slutt litt om oppgavens struktur og oppbygning. I del en redegjør jeg for naturalisme som kompatibilisme, motivasjonen bak en slik posisjon og hvorfor jeg tror den er fruktbar. Deretter tar jeg opp naturalisme som *metodisk* holdning, gjennom Jesse Prinz' begrep om empirisk filosofi. Til slutt i dette kapittelet ser jeg på hvordan *innholdet* i moralteorier påvirkes av empiri. Dette gjør jeg ved å bruke Owen Flanagans betraktninger rundt hva det innebærer å bedrive filosofisk moralteori som er psykologisk realistisk. I del to tar jeg opp debatten rundt dydsetikken. Jeg ser på dydsetikkens empiriske forpliktelser, Doris kritikk, og tar opp tre tilbakemeldinger på Doris' kritikk. Tilslutt i kapittelet vender jeg igjen blikket opp, og forsøker å se hva en slik "case" kan fortelle oss om hvilke utfordringer og muligheter som ligger i psykologisk realisme. I oppgavens tredje og siste del tar jeg opp en ny innvending mot dydsetikken. Den omhandler hvorvidt dydsetikerne legger *for* mye i sin naturalisme. Naturlige fakta avslører sin egen begrensning som fundament for teorier. Naturalismen lar det dermed bli igjen et stort felt for tenkning som ikke alene kan begrunnes i fakta om hvem vi er.

1. Naturalisme og kompatibilisme

Hvordan skal vi som moralfilosofor forholde oss til empiri? Står vi ovenfor et valg mellom og underorde etikken de ”hardere vitenskapene”, og adoptere ambisjonene som her finnes om å komme frem til fullkomment objektive og universelle lovmessigheter? Skal vi betrakte etikken som fullstendig isolert, og gi oss hen til en form for non-naturalisme hvor moralens prinsipper innehar en særegen ontologisk status? Eller skal vi simpelthen gi opp enhver ambisjon om teoribygging innenfor den filosofiske etikken? Peter Railton oppsummerer dilemmaet ved å liste opp en serie velkjente dikotomier fra det siste århundrets teoretisering rundt moral og etikk: ”Theory or anti-theory? Reductionist scientism or autonomist intuitionism? Naturalism or non-naturalism? Causes or reasons? Functions or meanings?” (Railton 2004: 268). Skal det langs disse linjene kjempes en kamp mellom etikk og vitenskap? Eller kan det være mer fruktbart å balansere mellom ytterpunktene, og plukke med seg poeng fra begge endene av spekteret? Kan vi finne et fundament for det normative innenfor et (natur)vitenskapelig verdensbilde, og samtidig bevare etikk og moral som meningsfulle og givende?

I ”An Experimental Philosophy Manifesto” (2008) beskriver Joshua Knobe og Shaun Nichols hvordan det har oppstått en ny interesse for forholdet mellom empiriske vitenskaper og filosofi i den senere tid. Filosofi av den typen som utføres fra lenestolen utfordres og suppleres av metoder som er mer eksperimentelt, naturvitenskapelig og empirisk orientert: ”It is true that *philosophical psychology* – the sort that can be done from an armchair and which is an assemblage of virtually every possible view of mind – is now giving way to *scientific psychology* ...” (Flanagan, Sarkissian og Wong 2008: 13). På sett og vis er ikke dette en ny måte å bedrive filosofi på. Før den analytiske filosofiens inntog på 1900-tallet var filosofer mindre interessert i å holde sin virksomhet adskilt fra andre disipliner, og mer interessert i betraktninger om hvordan verden henger sammen. Fakta kunne hentes fra ulike vitenskaper, så lenge de ble oppfattet som relevante. Ifølge Knobe og Nichols kan vi se at enkelte filosofer nå oppfordrer til en liknende innstilling. Den nye interessen skiller seg likevel fra den eldre tradisjonen på et viktig punkt, nemlig ved at fakta gjerne hentes fra konkrete empiriske studier (Knobe og Nichols 2008: 3). Jeg vil peke på to hendelser som kan belyse en slik bevegelse, nemlig hvordan man (gjen)oppdaget

psykologiens relevans for erkjennelsesteorien og dydsetikkens ”gjennoppstandelse” de siste femti årene. Jeg vil også vise at denne nye bevegelsen er en lovende utvikling.

I ”Naturalizing Ethics” trekker Owen Flanagan, Hagop Sarkissian og David Wong paralleller mellom naturalisme i erkjennelsesteorien og naturalisme i etikken ved å se på Quines naturalisme. I ”Epistemology Naturalized” (1969) argumenterer Quine for at erkjennelsesteori bør underordnes vitenskapelig psykologi. Empirismen har vist seg å ikke kunne gi oss et tilstrekkelig fundament for kunnskap. På bakgrunn av dette hevder han at erkjennelsesteorien som filosofisk disiplin bør oppgi sine forsøk på å bevise at vi kan ha kunnskap. I stedet bør man vie sin oppmerksomhet mot de psykologiske prosessene som faktisk ligger til grunn for våre oppfatninger av verden. Quine mener vi bør tenke på erkjennelsesteori som ”nok et kapittel i psykologien” (Flanagan, Sarkissian og Wong 2008: 11). Ved å gjøre erkjennelsesteori til psykologi konsentrerer man seg utelukkende om deskriptive spørsmål. Det er nok få som i dag vil dele Quines posisjon. Mange vil hevde at det er karakteristisk for erkjennelsesteori som filosofisk disiplin at man er opptatt av *normative* spørsmål: Om hvilke grunner vi har for våre påstander om kunnskap, og hvorvidt de prosessene som fører frem til kunnskap holder mål. Psykologien nøyer seg med å beskrive mentale prosesser og fenomener. Ved å gjøre erkjennelsesteorien underordnet psykologien mister man den normative komponenten. Man kan betvile hvorvidt Quines naturalistiske erkjennelsesteori er erkjennelsesteori i det hele tatt¹(Feldman, 2001).

Likevel, en interesse for forholdet mellom vitenskapelige data og filosofi ble vekket til live. Det må dessuten også bemerkes at nyere funn innenfor enkelte vitenskaper, slik som nevrovitenskap, biologisk psykologi og sosialpsykologi, har fremstått som svært relevante for filosofiske spørsmål. Man fikk dermed enda en grunn til å bry seg med empiriske vitenskaper. Når alt dette tas i betraktning velger mange å innta en form for mellomposisjon, hvor erkjennelsesteori er i kontakt med empirisk vitenskap, men uten å gjøres til en del av

¹ Quines svar på en slik innvending er interessant. Flanagan, Sarkissian og Wong plukker frem et sitat fra Quine: Ifølge han er naturalisering av erkjennelsesteori ikke det samme som å nøye seg med ukritiske beskrivelser av psykologiske prosesser. Erkjennelsesteori handler om å komme frem til de mest effektive midlene for å nå et mål, nemlig sannhet eller korrekte predikeringer. Det normative blir deskriptivt idet vi bokstaverer ut et endelig mål for virksomheten: ”For me normative epistemology is a branch of engineering. It is the technology of truth-seeking ...” (Quine sitert i Flanagan, Sarkissian og Wong 2008: 24). Arbeidet som må gjøres består kun i å komme med deskriptive redegjørelser for hvilke metoder som best mulig leder oss i mål. Hvis vi kunne fastslå et mål for moralen, slik som belønning i himmelen, ville den normative komponenten også her bli deskriptiv. Moral ville da omfatte de beste metodene for å oppnå belønning i himmelen.

disse vitenskapene. Flanagan, Sarkissian og Wong foreslår at en naturalistisk erkjennelsesteori kan tenkes å involvere to komponenter: en deskriptiv-genealogisk og en normativ komponent. Den deskriptiv-genealogiske komponenten vil inneholde informasjon om faktisk epistemisk virksomhet hentet fra et bredt spekter av forskningsfelter. Den normative komponenten samler sammen epistemiske normer som gir suksess. Dette er ikke å regne som enda en deskriptiv beskrivelse fordi disse normene innebærer en form for idealisering man er kommet frem til gjennom å skille ut de metodene som gir suksess fra de som ikke gir det. Vanlig menneskelig resonering følger, som vi alle altfor godt vet, ikke alltid de normene som gir suksess. En naturalistisk erkjennelsesteori er ikke *kun* psykologi. Likevel, de psykologiske reglene som den normative komponenten inneholder får sin normativitet fordi de er fruktbare i praksis. Å innføre naturalisme på denne måten innebærer å bevege seg mellom det deskriptive og det normative (Flanagan, Sarkissian og Wong 2008: 11).

Etter at erkjennelsesteorien har åpnet døren opp for empiriske vitenskaper, synes det som om turen nå har kommet til etikken. En rekke områder innenfor etikk ligger besnærende nær emner som behandles av empiriske vitenskaper. Temaer slik som vår psykologiske oppbygning, hvilke psykologiske elementer vi er i besittelse av, hvordan vi motiveres, hva som gjør at vi trives eller ikke trives, hvordan vi koordinerer samhandling og hvordan vi tilegner oss normer: Alt dette er emner som synes relevante for moral og etikk, og som behandles av empiriske vitenskaper. Utfordringen er at naturalismens problemer med å behandle normative spørsmål kanskje står enda mer sentralt i etikken enn i erkjennelsesteorien. I erkjennelsesteorien kan alle enes om ett mål for virksomheten, å finne ut hva som gjør kunnskap mulig, og man er relativt enig i hva denne kunnskapen er. I etikk og moral derimot er målene vi retter vår praksis etter mer diffuse. Vi kan muligens være enige i at moralen skal fremme velferd, men hva denne velferden består av, hvem den skal gjelde for, og hvilke normer som best fremmer den er mye mer kontroversielt. Moralens normer og regler vurderes derfor opp mot en standard som selv er åpen for vurdering. Å bringe moral og natur sammen synes dermed som en enda større utfordring her, enn i erkjennelsesteorien.

Kritikere av moralsk naturalisme påpeker derfor gjerne at alle mulige fakta om menneskelig moralsk praksis aldri vil kunne rettferdiggjøre denne praksisen. Ingen normative *bør* lar seg utlede av deskriptive *er*. Det er dog ikke sikkert at naturalisten behøver å gjøre seg skyldig i en slik feilslutning. Naturalisme trenger ikke å innebære annet enn å se på hva moralteorier

krever av verden, og avgjøre om verden er slik at teoriens påstander lar seg virkeliggjøre. Det handler med andre ord om å vurdere hvorvidt en moralteoris sentrale elementer er realiserbare i lys av empiri. Utfra en slik vurdering vil man kanskje kunne avsløre at noen moralteorier opererer med postulater om hvordan ting forholder seg i verden som er feilaktige: "scientific psychology ... may eliminate some of the classical views of mind on empirical grounds. If this happens, then our ethical theories will be framed by better background theories about our natures" (Flanagan, Sarkissian og Wong 2008: 13). Et slikt prosjekt er ikke det samme som å vurdere moralens endelige mål, eller konstruere nye mål utfra empiriske fakta. Vi kan heller betrakte det som å pålegge seg visse bånd når man bedriver moralfilosofi. Peter Railton kaller en slik holdning til forholdet mellom moralteori og empiri for *moralsk kompatibilisme*, hvilket innebærer å betrakte moralske kategorier og begrep på en slik måte at de lar seg forene, eller eksistere på en måte som er i overensstemmelse med empiriske vitenskaper (Railton 1995: 99). Kompatibilisten kan anerkjenne moralen som, i visse henseender autonom. Vi kan ikke hente noen normative *bør* ut av naturen alene. Samtidig forplikter han seg til å betrakte moralen på en slik måte at den ikke er i konflikt med våre beste empiriske teorier, "Reflective moral thought, if it is to avoid falling into systematic error or being saddled with 'necessary' illusions or unintelligible but necessary postulates must be underwritten by a compatibilist account of moral norms" (Railton 1995: 105).

For å igjen bruke naturalisme i erkjennelsesteorien som sammenlikningsgrunnlag: Til grunn for "naturaliseringen" av erkjennelsesteorien lå det ikke først og fremst en misnøye med tradisjonell erkjennelsesteori fordi den er feilaktig, men heller en tanke om at den er passé, eller representerer et utslitt paradigme (Sorell 1991: 3). Fremfor å konstruere teorier etter den tradisjonelle "lesten", fokuserer man på en ny metodikk hvor filosofi og andre vitenskaper er overlappende felter. Kan det tenkes at noe av det samme er tilfellet innenfor moralfilosofien? Det er mulig å betvile hvor fruktbare de klassiske problemene, slik som den naturalistiske feilslutningen, "The Open-Question Argument" eller debatten om internalisme versus eksternalisme, egentlig er. På sett og vis handler den "nye naturalismen" innenfor moralen ikke så mye om å argumentere mot eller tilbakevise disse innvendingene, som å bevege seg hinsides dem.

En moralsk kompatibilist forsøker på best mulig måte å forene det normative med det deskriptive. Railton presiserer at en slik forening ikke kan forseiles gjennom bevis som vil overbevise en moralsk skeptiker: "The link cannot, I believe, be effected by proof. It is no

more my aim to refute moral skepticism than it is the aim of contemporary epistemic naturalists to refute Cartesian skepticism” (Railton 1986: 163). Flanagan et al. motsetter seg også påstander om at å bedrive moralfilosofi innebærer å avdekke sammenhenger av logisk art. Naturalisten, ”... simply never claims to *demonstrate* anything. Why should he? Demonstration, Aristotle taught us long ago, is for the mathematical sciences, not for ethics” (Flanagan, Sarkissian og Wong 2008: 14). Moralens logiske autonomi vil dermed overleve naturalismen, og moralsk skeptisisme vil fremdeles være en koherent posisjon. Moralske naturalister vil likevel ikke la seg ”vippe av pinnen” av skeptikere eller fanebærere for moralens absolutte autonomi. De vil heller utfordre disse ved å peke på hvor lite fruktbare slike posisjoner er: ”... he presents to the skeptic a certain challenge, namely, to show how a skeptical account of our epistemic or moral practices could be as plausible, useful, or interesting as the account the naturalist ... (Railton 1986: 164). Kanskje kan man påstå at noe av det fruktbare ved naturalisme ligger i at de er ”dristige” i sin teoribygging. Med dette menes det at naturalisme er en produktiv posisjon fordi den forplikter seg til å være holdbar i flere dimensjoner. Naturalistiske teorier må *både* ta høyde for våre normative intuisjoner og bevege seg innenfor det som er empirisk holdbart. Railton peker på hvordan teoretisering generelt sett er mer produktivt når den ilegges begrensninger, eller må oppfylle kriterier, i flere sentrale dimensjoner. Dette er fordi slike begrensninger gir oss et mer lovende grunnlag for å kunne oppløse konflikter på teorinivå: ”One has such ample opportunities to be shown wrong or found unconvincing if one’s account must be responsive to empirical demands as well as normative intuitions” (Railton 1986: 205). Tilslutt kan vi kanskje betrakte naturalistens virksomhet som å ta sitt utgangspunkt i hva vi kan si om oss som mennesker, altså a posteriori, fremfor å ta utgangspunkt i moralske begrep eller intuisjoner og fra dette forsøke å utlede noe a priori.

Kompatibilisten kan, slik vi så, ikke fullstendig utelukke moralsk skeptisisme. Selv om vi kan vise til et nært forhold mellom moralske og naturlige egenskaper gjenstår et gap mellom dommer om empiri og dommer av moralsk art². Moralen beholder en logisk autonomi som tillater at man betviler hvorvidt moralens påbud virkelig er noe jeg *bør* gjøre (Railton 1995:

² Liknende gap finnes riktignok i andre disipliner også. Vitenskapens induktive metode innebærer også at valg mellom ulike teorier ikke fullt ut lar seg diktere av logikken, ”Morals are radically underdetermined by the merely descriptive, but so too, of course, are science and normative epistemology. All three are domains of inquiry where ampliative generalizations and underdetermined norms abound” (Flanagan, Sarkissian og Wong 2008: 14). Selv om vi ikke kan komme med vitenskapelige påstander med logisk sikkerhet er det svært få som påstår at man bør unnlate å drive med vitenskap, eller at vitenskap ikke finnes.

104). På den annen side vil naturalistens redegjørelse muligens kunne peke på noe som gjerne underkommuniseres i tradisjonell moralteori: Nærmest alle moralfilosofier kommer med deskriptive påstander. Moralen gir kanskje ikke mening for oss inntil den kobles opp mot en empirisk verdensorden. Både Railton og Flanagan et al. viser til hvordan sammenblanding av normativ holdbarhet og empirisk forklaring forekommer i tradisjonelle moralteorier. De fleste inneholder påstander om ”... philosophical psychology which postulate basic human dispositions that help or hinder morality (e.g. reason, emotion) and the sources of moral motivation. This is a ubiquitous feature of the moral tradition ...” (Flanagan, Sarkissian og Wong 2008: 13). Kantiansk moralteori kan betraktes som et forsøk på å forene sentrale normative intuisjoner med fakta om vår beskaffenhet som biologiske vesener, og aktører i en kausal verdensordning (Railton 1986: 206; Railton 1995: 97)³. Ofte vil teorier operere med moralske begrep som skal forklare aktørers handlinger. Fordi normative teorier gjerne er ment å ha et slik forklarende aspekt, kan man ikke uten videre avskrive forsøk på å betrakte disse teoriene i lys av verdens kausale strukturer som uholdbare (Railton 1995: 99).

I filosofenes teorier har moralen kanskje aldri vært helt autonom, og bør kanskje heller ikke være det. Gjennom å vise at moralen kan oppta en plass i en naturlig verdensordning kan naturalisten styrke vår tro på moralen som noe virkelig. Fremfor å undersøke hva moralen krever av verden, og konkludere med at verden ikke er konstruert på en slik måte at moralen kan oppta noen rettmessig plass i universet, vil en kompatibilist også spørre: Hva krever verden av moralen? Hva er det som gjør at noen fakta motiverer oss? Eller at noen ting oppleves som bærere av en verdi som ikke er å betrakte som kun instrumentell? En naturalistisk redegjørelse for moralen gir oss grunn til å tro at våre antagelser og intuisjoner om forholdet mellom moral og natur ikke er illusjoner. Å være kompatibilist innebærer å sammenkoble moral og natur uten å ende i scientisme eller andre posisjoner som er undergravende eller truer moralens egenart: ”Showing compatibility is a way of promoting the autonomy of moral reasoning, for it would show that we are not running afoul of our own convictions about the relation of the moral to the natural” (Railton 1995: 105). Filosofer som i liten grad er opptatt av å integrere moralen med resten av verden risikerer å ende opp med teorier som er for abstrakte eller legger moralens kriterier utenfor aktørenes rekkevidde.

³ Og en naturalist vil anklage Kant for å tillegge visse moralske intuisjoner så stor vekt at det skader teoriens empiriske holdbarhet.

Teorier som postulerer overmenneskelige evner til moralsk resonering, evner som ikke kan virkeliggjøres i noen aktører overhodet, risikerer å miste sin relevans for oss. Man pleier vanligvis å hevde at *bør* medfører *kan*, men hvis jeg prinsipielt ikke *kan*, hvilken grunn har jeg da til og i det hele tatt bry meg med *bør*? Kanskje kan det normative bare komme til sin rett hvis vi integrerer moral og natur?

Dette fører oss til moralsk motivasjon. Naturalister forsøker å bringe moralsk motivasjon ”tilbake på landjorden”. Moralsk motivasjon kan betraktes som en psykologisk mekanisme, på lik linje med andre psykologiske fenomener:

It is largely a matter of psychological, sociological, and anthropological inquiry why different sorts of things are motivating reasons, that is, why certain reasons and not others motivate at different times and places. The role for the normative naturalist is to recommend ways of finding *good* reasons for belief and action and to indicate why it makes sense to be motivated by such reasons (Flanagan, Sarkissian og Wong 2008: 7)

Fakta om hva som motiverer oss i en hverdagslig kontekst vil kanskje ikke kunne overbevise en som først og fremst er interessert moralske dommers metafysiske og ontologiske status, men på grunnlag av dette å avskrive psykologiske data som irrelevante for moralteori er en forhastet konklusjon. Om ikke annet bør psykologien minne oss på at moral er praktisk filosofi. Moralen handler om å komme opp med idealer, og man vil anta at disse idealene er ment å kunne i en viss forstand virkeliggjøres. På denne måten er moralen alltid i kontakt med verden, og særlig via de psykologiske strukturene som gjør at moralens idealer kan virkeliggjøres. I filosofisk moralpsykologi fremmes det dermed en rekke påstander om eksistensen av ulike psykologiske elementer, og om hvordan disse forholder seg til hverandre. Se for eksempel på følgende utsagn om nonkognitivismen om moralske dommer:

Psychologically, since making a moral judgement requires our having a certain desire, and no recognition of a fact about the world could rationally compel us to have one desire rather than another, our judgement must really simply be an expression of that desire, or perhaps a complicated disposition to have that desire. (Smith 1994: 10)

Her trekkes det substansielle filosofiske slutninger på grunnlag av noe som vi må anta er psykologiske fakta. Denne måten å betrakte forholdet mellom moralske dommer, motivasjon og psykologiske elementer slik som trosoppfatninger og ønsker, er knyttet opp mot en humeansk modell for moralsk motivasjon. Humes modell har relativt få bestanddeler, og

kondenserer elementene som trengs for å redegjøre for moralsk praksis ned til "the bare essentials". Dette har selvsagt sine fordeler hva angår argumentatorisk klarhet, og fremhever filosofiske distinksjoner som er av stor betydning, slik som forskjellen mellom grunner og årsaker. Men det krever også varsomhet. Prisen man betaler for slik klarhet synes å være at man risikerer å gjøre moralen til noe mer teknisk enn hva det faktisk er, eller forventer at moralske grunner er av en helt annen art enn andre typer grunner. Dette bringer meg til dydsetikken, slik jeg annonserte i begynnelsen av kapittelet. Dydsetikeren Rosalind Hursthouse poengterer at den moderne moralfilosofien er svært betatt av begreper som rasjonalitet og plikt. Dette gjør at de forventer at når man setter adjektivet "moralisk" foran et begrep skal dette utløse en egen, magisk, og rettferdiggjørende kraft (Hursthouse 1999: 122). Noe av drivkraften bak en humeansk modell er at man ved hjelp av den kan rydde i verdens metafysiske bestanddeler ved å avklare hva moralske fakta egentlig er. Kanskje kan man spørre seg om hvorvidt visse metaetiske foretak ikke så mye handler om å gi et adekvat bilde av moralen, som å skape "metafysisk orden". Ifølge Julia Annas er moralsk kunnskap blant annet praktisk kunnskap, et faktum mange moderne filosofer ikke tar konsekvensene av: " ... it is significant that so many modern positions in moral epistemology not only refuse to start here, as we prephilosophically do, but make the practicality of moral knowledge mysterious, or even regard it as forcing us to accept debunking conclusions about value" (Annas 2001: 254). Ifølge Annas har moderne moralfilosofi vært for opptatt av å finne plass til moralen i et system som kanskje ikke er egnet til romme denne typen kunnskap. Dermed fremhever dydsetikere gjerne en mangfoldighet i vår moralpsykologi: de er opptatt av rasjonalitet, men ikke kun som instrumentell, emosjoner som kognitive, fremfor som ikke-rasjonelle drifter, og moralsk motivasjon som noe som opptrer i ulike grader. En slik mangfoldighet gjenspeiles i de elementene dydsetikerne regner som moralsk relevante. De velger å betrakte hele aktøren, fremfor kun å konsentrere seg om et lite utsnitt. Slik sett er dydsetikken på linje med nyere teorier innen kognitiv vitenskap som fremmer et liknende helhetsperspektiv i motsetning til de noe eldre teoriene som gjerne tenkte på bevisstheten som en datamaskin. Slik Mark Johnson sier om nyere teorier innen kognitiv vitenskap...

... recognizes the embodied and imaginative character of all human conceptualization and reasoning. It focuses on the social, interactive, evolutionary character of human experience and understanding. It looks at human cognition as embodied in a growing biological organism that is interacting and co-evolving with its physical, social, and moral environments" (Johnson 1996: 51)

Dermed kan det se ut til at en måte vitenskapen forandrer moralteorier er ved å kreve en mer mangfoldig moralpsykologi. Kanskje er menneskesinnet et område som ikke tillater en streng sparsommelighet i katalogiseringen av bestanddeler. Det ser heller ikke ut til at samspillet mellom disse bestanddelene følger logikkens regler. Påvirkning fra trosoppfatning til ønske skjer kanskje ikke på en måte som er absolutt i samsvar med hva fornuften eller logikken påkrever. Mekanismene kan være, slik Railton beskriver det, *bare* psykologiske (Railton 1986: 178). For naturalisten er de likevel av betydning fordi han vil bygge teorier hvor slike grunnleggende prosesser tas på alvor, og er råmateriale for hva vi kan si om oss som moralske vesener.

Til nå har vi betraktet et tettere forhold mellom moralfilosofi og vitenskap som et ideal det er verdt å strebe etter. Å være empirisk sensitiv kan gi opphav til moralteorier som er mer gjenkjennbare, realistiske og revitaliserte. Det finnes selvsagt berettigede bekymringer knyttet til et slik ideal. At et tettere forhold mellom moralfilosofi og vitenskap skal virke undergravende for moralen er en reell trussel. Som vi skal se er det noe foruroligende med visse psykologiske studier som viser at motivene vi handler på ikke alltid er slik vi tror. De kan være av en type vi ikke er i stand til å påvirke, eller ikke så edle som vi liker å tro. Senere i oppgaven skal vi se at studier viser at i alminnelige situasjoner kan grunnene vi handler på være ukjente for oss, og av en art vi vanligvis ikke ønsker å identifisere oss med. Dette er kanskje først og fremst et praktisk problem: Det gjør noe med mitt selvilde som moralsk aktør at jeg så å si ikke kan stole på meg selv og min egen handlekraft. Noen illusjoner brister idet vi forstår at mange filosofiske modellens høye krav til kunnskap om, og tilgang til selvet ikke lar seg virkeliggjøres i aktører. Slike forstyrrende fakta blir ikke mindre farlige ved at vi rømmer inn i filosofiske modeller hvor ulike grunner fint lar seg dele inn i ulike kategorier. Det er et spennende tema hvorvidt man kan lage substansielle moralteorier som tar høyde for denne typen utfordringer.

For å samle noen tråder: Vi har sett hvordan en kontakt mellom filosofi og empiriske vitenskaper kan gi opphav til moralsk kompatibilisme. Denne posisjonen innebærer å tilpasse moralteorier etter kunnskap vi har om verden, og da særlig kunnskap produsert av empiriske vitenskaper. En slik tilpassning vil ikke skje med logisk nødvendighet, men fordi de fleste moralteorier vil inneholde deskriptive elementer er en slik tilpassning likevel å betrakte som et gode. Teorier som tar konsekvensen av hva vitenskapen kan fortelle om hvem vi er vil være mer gjenkjennbare, og derfor kunne operere med et normativt *bør* som synes mer relevant for oss. Den siste betraktningen jeg gjør meg mot slutten av delkapittelet peker

riktignok mot en utfordring: Vitenskapen vil kunne komme opp med informasjon om oss som ikke er av en oppbyggelig art. Hva kompatibilister kan gjøre med slik kunnskap gjenstår å se. Videre i dette kapittelet vil jeg se på hva moralsk kompatibilisme fører med seg av metodiske og teoretiske forpliktelser. Jesse Prinz er opptatt av hvordan filosofi og vitenskap ikke er to adskilte sfærer, men faktisk befinner seg langs samme spekter, og at dette får konsekvenser for hvordan man behandler filosofiske og normative spørsmål. Han beskriver hvordan man nå finner filosofer som er mer praktisk rettet i sitt forhold til andre vitenskaper. De går aktivt ut og benytter seg av empiri fra andre felter i sin egen forskning. Hos Owen Flanagan finner vi en moralsk naturalisme innenfor ett spesifikt område av moralfilosofien, nemlig moralpsykologien. Han inntar en psykologisk realistisk posisjon, og mener at overensstemmelse med vitenskapens empiri bør være et relevant kriterium når man vurderer og konstruerer teorier om moral.

1.1 Prinz om empirisk filosofi

Vi har sett at det er gode grunner til å mene at moralen både kan og bør være i samsvar med det vi vet om resten av verden. Dette er kanskje ingen genuint ny påstand, men i den senere tid har denne erkjennelsen produsert en ny måte å tilnærme seg filosofiske emner på. Et naturlig spørsmål å stille en moralsk kompatibilist er: Hvis man ønsker å bedrive moralfilosofi i overensstemmelse med de krav empirien stiller, hvordan vil en slik filosofi se ut? Det Jesse Prinz kaller *empirisk filosofi* gir et bilde av hvordan man kan stille seg til bruk av vitenskapelige data i filosofien. Han forsvarer en slik metode gjennom å vise at tradisjonelle metoder innenfor filosofien ikke opererer hinsides empiriske sfærer. Skillene mellom filosofi og ulike empiriske vitenskaper ikke er absolutte, men mer å regne som ulike deler av ett stort lappeteppe. Enkelte filosofers empiriske oppvåkning er dermed et nødvendig gode, og en naturlig konsekvens når et kunstig skille oppløses.

En empirisk disiplin bygger i stor grad på data fra observasjon, og da særlig observasjoner som kan gjentas og deles, og fungere som elementer i teorier om hvordan ting forholder seg i verden (Prinz 2008: 190). Filosofi regnes gjerne for ikke å være empirisk på denne måten. I “Empirical Philosophy and Experimental Philosophy” (2008) vil Prinz vise at dette er en naiv og feilaktig holdning. I likhet med empiriske vitenskaper hviler filosofien på en metode som innebærer en form for observasjon. Hva er det som tradisjonelt sett anses å være

filosofiens metode? Prinz mener at en sentral del av filosofiens arbeid kan karakteriseres som begrepsanalyse. Fra lenestolen gjennomfører filosofer tankeeksperimenter, eller andre former for introspeksjon, som skal plukke ut visse intuisjoner. Disse intuisjonene er ment å gjenspeile konseptuelle sannheter. Ifølge Prinz er dette en form for observasjon, dog uten bruk av (de ytre) sansene eller måleverktøy. Introspeksjon er ikke av natur fundamentalt ulik den type observasjon som empiriske vitenskaper baserer seg på. Det som observeres, altså våre intuisjoner, kan deles og sammenlignes med andres, og vi kan granske den samme intuisjonen flere ganger. Dessuten er fleste begreper oppstått i samspill mellom ytre stimuli, og introspeksjon er således en form indirekte betraktning av den ytre verden. Det er feilaktig å tro at introspeksjon gir oss tilgang til en sfære av medfødte intuisjoner eller noe à la Platons ideer, ”In sum, armchair conceptual analysis can be characterized as an introspective memory retrieval process. As such, it can be regarded as a form of observation” (Prinz 2008: 191). Vi vil kanskje vegre oss for å kalle begrepsanalyse en empirisk metode. Uavhengig av hvordan vi velger å klassifisere en slik metode er Prinz’ poeng at begrepsanalyse ikke er fri for empiri. Forskjellen mellom introspeksjon som observasjon og det som tradisjonelt sett har blitt ansett som empirisk observasjon er ikke absolutt. Begrepsanalyse er ikke en rent a priori virksomhet.

Det er dessuten en rekke områder hvor introspeksjon kommer til kort. Å benytte seg av kun introspeksjon når man vil undersøke begreper er kanskje ikke ”god metode”. Introspeksjon er sårbar for feiltagelser fordi vi lar oss påvirke av forventinger, normer og teorier⁴.

Dessuten er intuisjonene introspeksjonen fanger opp ikke nødvendigvis av allmenn art, men er ulike gitt visse variabler, slik som kultur og kjønn. Tradisjonelt sett har filosofer vært for lite sensitive for variasjon i intuisjoner, mener Prinz: ”In philosophy, debates often collapse into intuition mongering because defenders of opposing views are equally confident about conflicting intuitions” (Prinz 2008: 191). Stilt overfor slike konflikter bør man erkjenne at intuisjoner ikke eksisterer i en lukket sfære, men påvirkes av en rekke faktorer. I tillegg til påvirkning utenfra, er det strukturelle aspekter ved bevisstheten som gjør at den på egen hånd produserer forstyrrende elementer. Hukommelsen er konstruktiv, og vi kan lure oss selv ved å konstruere fantasifostre. Vår bevissthet er dessuten ikke gjennomsiktig, og det

⁴ Prinz har kanskje rett i å peke på dette som en begrensning ved introspeksjon, men det er nærliggende å anta at dette er en kilde til feil i vi finner i alle former for observasjon. Forventninger og bakgrunnskunnskap er elementer som påvirker våre observasjoner generelt sett.

kan finnes aspekter ved våre begreper som er utilgjengelige for oss. Fordi begreper lagres som prototyper⁵, mener Prinz, kan de ikke vise til noe a priori. De er heller å betrakte som konstruksjoner basert på erfaring og innsamlet materiale. Dette betyr at introspeksjon hverken er en ufeilbarlig metode eller gir innpass til sannheter som er a prioriske eller universelle. Derfor bør vi være varsomme når vi benytter oss av introspeksjon og appellerer til intuisjoner, og også supplere med andre metoder når vi ønsker å redegjøre for begreper (Prinz 2008: 191-192). Det er nettopp dette empiriske filosofer gjør. Begrepsanalyse forekommer ikke bare innenfor filosofien, også innenfor mer empirisk orienterte vitenskaper finner vi semantiske analyser. Her finnes det mye godt materiale for filosofer å gripe fatt i. En filosof som utelukkende gjør sin forskning fra lenestolen undersøker kanskje mer seg selv og sin kulturs fordommer, fremfor å avdekke absolutte og universelle strukturer. En slik begrepsanalyse å betrakte som, i verste fall dårlig utført, deskriptiv sosialpsykologi (Flanagan 1991: 28).

For å kunne bruke intuisjoner som grunnlag for teori må man kunne si noe om kilden til denne intuisjonen, og sannsynliggjøre at denne kilden er pålitelig. Man bør bruke eksperimentelle data for å forstå prosessene som ligger til grunn for enkeltintuisjoner, og i lys av dette vurdere gyldigheten til disse intuisjonene. Hvilke kriterier vi legger til grunn for gyldighet når vi vurderer intuisjoner kan selvsagt ikke gis oss av empiri. Vi står likevel ikke uten noe vurderingsgrunnlag. Vi har visse antagelser om hvilke intuisjoner og hvilke prosesser som er akseptable. Vi kan revurdere og justere disse antagelsene i lys av empirien. Gjennom en slik justering kan vi komme frem til en bedre forståelse av intuisjoner og prosessene som genererer dem. (Knobe og Nichols 2008: 10-11).

Er det noe med normative spørsmål som gjør at de utelukkende krever filosofisk behandling? Nei, mener Prinz. Filosofer har ingen særlig autoritet i spørsmål angående normativitet. Filosofien kan, gjennom begrepsanalyse, kaste lys på sentrale normative begreper, slik som frihet, lykke eller bør-utsagn. Ofte er det viktig med bergrepsavklaring for å håndtere normative påstander. Hva innebærer det at noe er en plikt? Hva innebærer frihet og er frihet en nødvendig betingelse for moralsk ansvar? Dette er sentrale spørsmål knyttet opp mot moralske begreper som krever nærmere granskning. Slik vi allerede har sett er begrepsanalyse ingen lukket filosofisk virksomhet. Psykologer, sosiologer og

⁵ Begreper som prototyper vil jeg komme tilbake til i kapittel 2.1

antropologer kan bidra med å bestemme hva begreper egentlig inneholder, og gjennom empirisk forskning i disse feltene kan man øke forståelsen av hva det er vi refererer til når vi sier at en handling er noe vi *bør* gjøre eller at noe er en plikt. Sett at empirisk forskning avdekker at menneskers moralske dommer produseres gjennom et samarbeid mellom emosjoner og rasjonalitet. Dette ville være et funn som er filosofisk interessant. Det sier noe om fundamentet for moral hos folk flest, og fordi filosofer også er mennesker sier det kanskje også noe betydningsfullt om fundamentet til moralteorier. Filosofiens eget distinkte bidrag, og det som gjør at (empirisk) filosofi⁶ ikke vil kollapse inn i disse mer empirisk orienterte feltene, er at filosofer kan tillate seg å gjøre generaliseringer, komme opp med argumenter, bygge teorier, og så å si bringe empiriske resultater et steg videre. Det er dermed ikke sagt at det kun er filosofer som gjør dette, det finnes psykologer som er filosofisk anlagt. Prinz antyder at adskillelsen er et resultat av de ulike disiplinenes praksis: Eksperimentelle psykologer er gode til å samle empiri og utføre studier, filosofer er gode på å være teoretiske:

Empirical philosophers are theoreticians. They are not collecting data, but rather constructing arguments and theories, trying to figure out the implications of empirical studies. Empirical philosophers often integrate a range of different studies, from different branches of science ... (Prinz 2008: 205)

Selv om feltene overlapper hverandre, er forskjellene mellom de ulike disiplinene ofte relativt klare. Slik Prinz beskriver det: det er sjeldent vi har vansker med å avgjøre om en artikkel vi leser skal regnes som filosofi eller psykologi (Prinz 2008: 206). Men forskjellene er også kontingente og ikke å regne som absolutte. Eksperimentelle psykologer må noen ganger vende seg mot lenestolen å reflektere over teori og metode, og filosofer kan i enkelte tilfeller selv utføre empiriske studier⁷.

Frem til nå har vi sett på moralfilosofien som deskriptiv virksomhet, men handler ikke moralfilosofi vel så mye om hvordan ting *bør* være, fremfor hvordan ting *er*? Ja, vil Prinz

⁶ Prinz bemerker at ikke alle filosofiske disipliner vil ha like mye nytte av empirisk forskning. Erkjennelsesteori og etikk synes å ha mye med empiri å gjøre, men hvorvidt empiri vil ha noe særlig relevans for metafysikk er tvilsomt (Prinz 2008: 207).

⁷ Prinz gir navnet *eksperimentelle filosofer* til de som skaffer seg førstehånds empiri, fremfor empiriske filosofer som systematiserer og teoretiserer over allerede eksisterende forskning (Prinz 2008: 196).

påstå. Det er derimot feilaktig å tro at filosofer er i en privilegert posisjon til å avgjøre hvilke påbud som fortjener vår lydighet, og hvilke måter å ordne sosialt samvær på som er moralsk sett best. Filosofien kan komme med svært opplysende innsikter om hva vi mener når vi kommer med slike normative påstander, men dette er å regne som et deskriptivt prosjekt, noe filosofien ikke er alene om. Etter at vi har fiksert moralens begreper er arenaen åpen for alle som mestrer disse begrepene, ”So the distinctively normative character of moral discourse is not the aspect of moral discourse that makes morality seem more philosophical than scientific” (Prinz 2008: 195). Fremskritt innen etikken er ikke å regne som kun intellektuell utvikling, men er også rettet mot hva vil alle er i stand til å motiveres av. Moral og etikk er praktisk orienterte emner hvor det er viktig å kunne generere handling. Derfor kan ikke etikken gjøre det den er ment å gjøre hvis den kun behandles av en liten gruppe eksperter. Moral og etikk er et prosjekt som angår oss alle: ”Because the questions are about how men should live and how societies should be arranged, the answers must be accepted and internalized by many people to be effective” (Nagel 1980: 199).

Det er viktig å huske at Prinz ikke forsøker å gjøre filosofien overflødig. Hans argumenter er først og fremst rettet mot de som mener at filosofi best gjennomføres isolert fra andre vitenskaper. Det er kanskje optimalt å diskutere normative spørsmål innenfor filosofien fordi man her mestrer og har god kjennskap til sentrale moralske begrep. Normer for nøyaktighet og klarhet gjør at filosofien kanskje er en særdeles velegnet arena for diskusjoner av denne typen. Risikoen for å komme med feilaktige eller uholdbare påstander er dermed mindre. I lys av Prinz’ redegjørelse ser vi at filosofien ikke har noen enerett på normative spørsmål. Det er likevel grunn til å hevde at filosofiens egenart gjør at den er godt egnet til å behandle etikk og moral, men dette gjelder kun så lenge filosofien er åpen for innspill og korrigeringer fra empiriske vitenskaper. Uten en slik åpenhet er alle filosofiens påbud om klarhet og nøyaktighet uvirksomme: Vi vil likevel risikere å gjøre grove feil eller konstruere vedvarende illusjoner.

1.2 Flanagan om minimal psykologisk realisme

Jeg vil fremheve to punkter som synes for meg å være sentrale i forholdet mellom kompatibilistens moralteori og empiri: Først, empiri er *korrigerende* for teorier. Det vil si, empiri kan avsløre filosofiske fordommer som er bakt inn i moralteorier. Dette vil blant

annet innebære, slik vi så hos Prinz, å peke på at intuisjonene en teori bygger på ikke er fundamentale (i den forstand at de ikke trenger begrunnelse). Intuisjoner alene er ikke nok til å konstruere en gyldig moral (man må i så fall argumentere for at intuisjonene er av en bestemt art som gjør at de kan innlemmes i en teori. Det i seg selv vil sannsynligvis innebære empirisk forskning). Mitt andre punkt går ut på at empiri gjør moralen *engasjerende*. Med dette mener jeg at gjennom å konstruere moralteorier som er i samsvar med våre beste teorier fra empiriske vitenskaper sikrer man i større grad moralens normative grep om oss. Et slikt grep kommer av at teorien oppfattes som å tegne et gjenkjennbart og praktisk relevant bilde av moralske aktører og sosial samhandling. Spesielt viktig for teoriens normative elementer er moralsk motivasjon. Min påstand er at for å kunne gi en tilfredsstillende redegjørelse for motivasjon er samsvar mellom moralteori og psykologi viktig. Moralens relevans som motivasjonsfaktor økes gjennom å tegne et bilde av den moralske aktøren med en psykologisk struktur vi kan kjenne oss igjen i.

Owen Flanagan tar opp forholdet mellom moral og psykologi i *Varieties of Moral Personality* (1991). Også han mener at empiri, og her da særlig empiri fra psykologien, bør virke korrigerende på filosofien: "... scientific psychology has the potential for destabilizing, as well as for developing and refining, certain assumptions underlying traditional moral theory ..." (Flanagan 1991: 21). Flanagan ser deretter nærmere på hvordan en økt bevissthet om psykologiske fakta kan forme moralteorier. Her fremheves det at å betrakte psykologi som relevant innebærer at man må forholde seg til spørsmål rundt personlighet, om hva som er gode liv, og hva som er velfungerende mennesker. Psykologien utvider altså moralen til å ikke bare omfatte normer for sosial interaksjon, men også intrapersonelle aspekter som ikke har noen åpenbar nytteverdi for koordineringen av interpersonelle forhold (Flanagan 1991: 19). En annen måte psykologien virker utvidende på er gjennom å vise hvordan en streng adskillelse mellom moralsk overveielse og andre mentale fenomener ikke er empirisk plausibelt. I logikk og teoretisk tenkning lar kanskje moral seg klart adskille fra alt annet. Menneskesinnet er mer pragmatisk anlagt, og tillater at en rekke ulike elementer av bevisstheten samarbeider om å felle moralske dommer (Flanagan 1991: 19). En bevissthet om psykologiske fakta kan dermed gi utslag i en redegjørelse som er mer omfattende med henhold til moralsk motivasjon, og hva mennesker oppfatter som viktig og verdifullt. Dette kan i sin tur resultere i en moral som er mer levende og motiverende.

Flanagan skriver at de fleste teorier om moral streber etter på en eller annen måte å være i overensstemmelse med menneskelig psykologi. De fleste filosofer regner vår psykologiske oppbygning som relevant i en eller annen grad. Hvis en teori konstruerer moralsk gode aktører som umulig er menneskelig realiserbare vil de fleste regne dette som en alvorlig svakhet ved teorien (Flanagan 1991: 26). Det er i det hele tatt vanskelig å forestille seg en teori som ikke streber etter å være i samsvar med virkeligheten på en eller annen måte. Derimot er det enkelt nok å komme med eksempler på teorier som kan anklages for å mislykkes i å opprette et slikt forhold til verden. Konsekvensialisme, for eksempel, kan anklages for å kreve en umenneskelig kapasitet til å kalkulere og vurdere mulige handlingsutfall. Kriteriet, som altså de aller fleste teorier streber etter å være i overensstemmelse med, oppsummerer Flanagan i et prinsipp om minimal psykologisk realisme (PMPR):

Make sure when constructing a moral theory or projecting a moral ideal that the character, decision processing, and behavior prescribed are possible, or are perceived to be possible, for creatures like us (Flanagan 1991: 32)

Som tidligere nevnt er dette et prinsipp som fanger opp noe de fleste teorier aspirerer etter, men det gir oss også et kriterium for å vurdere hvor godt teoriene lever opp til idealet de sikter mot: ”PMPR is meant to be both descriptive and prescriptive. It picks out an aspiration of almost all moral theories, and it sets out a criterion for evaluating theories in terms of this aspiration” (Flanagan 1991: 33). Et sentralt poeng for Flanagan er at innføringen av et slikt påbud ikke vil kunne diktere hvordan teorier skal se ut. Prinsippet gir stort spillerom for hva man kan tillate seg med henhold til moralteori og moralpsykologi. Selv om å innføre et slikt prinsipp muligens utelukker noen måter å teoretisere om moral på, er det, som Flanagan påpeker, ikke tilstrekkelig til å plukke ut én teori fra alle de andre. Til det er prinsippet for inkluderende. Det må poengteres at slik Flanagan formulerer PMPR er det bare et krav at den psykologiske strukturen *oppfattes* som mulig. Det medfører at en teori kan operere med en struktur som ikke er faktisk realiserbar, men som er såpass gjenkjennelig for oss at den kan fungere som regulativt ideal. Dydsetikken kan for eksempel leses som å gi et slikt regulativt ideal.

Flanagan opererer med begrepet *moralsk personlighet*, et begrep som er ment å omfatte de personlighetsstrukturene og egenskapene som er moralsk relevante, i den forstand at de tar del i prosesser som igangsettes når vi handler i hensyn til det moralske. Vi så ovenfor at

tradisjonelt sett har filosofer vært relativt sparsommelige i sin omgang med personlighetsstrukturer: kun et lite antall elementer regnes som relevante. Å konstruere teorier som utviser en slik enkelhet kan ofte være et gode, men kun innenfor områder som tillater det. Moralen er kanskje et felt hvor teorier med få elementer ikke passer seg. Innledningsvis henviser Flanagan til hvordan virkelige moralske idealer, slik som Moder Teresa og Gandhi, ikke synes å passe inn i tradisjonelle filosofiske teorier om hva som konstituerer genuin moralsk motivasjon, selv om deres handlinger godt kan være i overensstemmelse med teoriens påstander om hva som er rett. Handlingene til virkelighetens idealer synes å være komplekse og lite systematiserte i sin opprinnelse:

Any actual human being will possess numerous personal characteristics, some moral, some nonmoral, that he or she displays in moral life. These characteristics will color and give unique character to any principles he or she lives by, and they may often work independently of, but not necessarily inconsistent with, such principles. What Mother Theresa does accords with both the categorical imperative and the principle of utility, but she is hardly guided by either. (Flanagan 1991: 7)

De moralske idealene den tradisjonelle moralfilosofien serverer dreier seg enten om aktører som kalkulerer handling ut fra et overordnet og allment prinsipp, eller som innehavere av et fullkomment sett med dyder. Problemet, mener Flanagan, er at filosofiens idealer ikke kan gjenfinnes i noen virkelige idealer. Dessuten, han peker her særskilt på dydene, er ofte sentrale teoretiske elementer ikke koherente eller er begreper vi ikke forstår tilstrekkelig (Flanagan 1991: 11). Flanagan er opptatt av at moralske personligheter kan realiseres på mange ulike måter, og å rette seg etter PMPR vil kanskje innebære å forlate de tradisjonelle teoriens sparsomme rasjonering av prinsipper. Å være psykologisk realistisk i denne sammenhengen, håper Flanagan vil resultere i at "... some of the many and rich varieties of moral personality become more visible, and ... the ideal of moral character as constituted by a single general-purpose principle or an ideally configured complement of virtue loses plausibility" (Flanagan 1991: 12). Hvis vi følger Flanagan her, vil PMPR lede oss i en slik retning?

Det er et faktum at teorier om moral utarbeides av mennesker, og virkeliggjøres via enkeltmenneskers psykologiske strukturer. Det er dermed svært lite trolig at noen teorier vil være fullstendig urealistiske i den forstand at de er utenkelige eller ikke innehar noen aspekter som er gjenkjennbare for oss. Psykologi vil ikke oppfattes som begrensende for

teoribygging på en slik måte: dette er en begrensning filosofene fint klarer å pålegge seg selv uavhengig av hvor relevante man oppfatter psykologiske fakta å være. Det psykologien derimot kan bidra med er å avgjøre hvor urealistisk en teori er. Det vil si, i hvilken grad den lar seg virkeliggjøre, og dessuten, om den pålegger enkeltaktører krav som er for sterke. I virkeligheten, altså rent deskriptivt, oppstår ulike moralske personligheter som produkt av egenskaper som er ilagt oss fra naturens side, og også som et resultat av ulike sosiale, økonomiske og statsmessige konstruksjoner. Gitt at moralske personligheter oppstår i samspill med materiale fra begge disse dimensjonene, og at den sosiale dimensjonen rommer svært mange ulike variasjoner, kan det gi opphav til et stort antall mulige moralske personligheter. Et samfunn med stor sosial integrasjon vil for eksempel kunne gi opphav til en moralpsykologi hvor den enkeltes prosjekter og ambisjoner lett lar seg overkjøre av hensyn til fellesskapets beste. Mens mer liberale samfunn kanskje vil gi opphav til personligheter hvor det å ofre egne livsprosjekter på samme måte nærmest oppfattes som en umulighet. Å følge en moral som oppfordrer til å oppgi vårt personlige perspektiv i stor grad er kanskje svært lite realiserbart for de med en vestlig bakgrunn, men det er ikke en umulighet innenfor andre kulturer, for eksempel buddhistiske. Siden samfunnet kan organiseres på så mange ulike måter må vi også anta at det samme gjelder moralske personligheter.

Et skille det er verdt å merke seg er skillet mellom teorier som gir kriterier for hva som er rett (uten hensyn til menneskelig psykologi), og teorier som redegjør for prosedyrer som leder frem til rett handling (Flanagan 1991: 34). Kun teorier av den sistnevnte typen som risikerer å være i uoverensstemmelse med PMPR⁸. Problemet er at det er svært få, om noen, teorier som kun begrenser seg til å gi kriterier for det gode eller rette. De aller fleste teorier vil også forsøke å gi en redegjørelse for moralsk motivasjon, altså hvordan en aktør må gå frem for å kunne oppfylle de kriteriene som gis for rett handling. Dette innebærer å komme med påstander av moralpsykologisk art. Etikk handler om hvordan verden *bør* være, om hvordan vi best mulig kan organisere menneskelig samhandling, og om å styre adferd, ikke bare trosoppfatninger (Nagel 1980: 198). Flanagan uttrykker det slik: det handler om å sette standarder for menneskelig adferd og karakter, for hva som er verdt å strebe etter. Forstår

⁸ For å uttrykke det noe klarere: Prinsipielt sett trenger ikke en teori av førstnevnte type å forholde seg til spørsmål om hvorvidt det gode eller rette er realiserbart for vesener som oss. Det er en reell mulighet at det gode befinner seg hinsides hva vi er i stand til å gjøre. Rent praktisk derimot vil nok de fleste også på dette nivået forholde seg til hva som er mulig gitt en menneskelig psykologi.

man det slik er forholdet mellom moral og sosial praksis asymmetrisk, moralen gir kriterier for den sosiale praksis (Flanagan 1991: 24). Verden slik den er har ikke noen innvirkning på hvordan man konstruerer det normative. Slik kan man argumentere for fullstendig isolering av etikken fra naturen. For å bringe oss tilbake til utgangspunktet, nemlig forholdet mellom kriterier for rett handling og motivasjon, kan vi begynne med å merke oss at kriteriene for det gode eller rette fremmes som kriterier som verden bør formes i overensstemmelse med. Målet er at den virkelige verden skal sammenfalle med en idealisert verden, men dette kan kun forekomme ved at noe endres i den virkelige verden. Her må moralsk motivasjon bringes inn. Enhver seriøs moralteori må tegne et bilde av hva slags motivasjonell struktur som kreves for at det teorien fremmer som godt eller rett skal kunne virkeliggjøres. Videre kreves det argumenter for hvorfor vi skal tro at en slik struktur er mulig. Vi motiveres av en moral vi kan kjenne oss igjen i, og som fremstår som noe verdt vår innsats og beundring, "[it] will draw on an image of ourselves we are capable of admiring and to which we can in some sense imagine conforming" (Flanagan 1991: 26). Moralen får sin relevans gjennom å være motiverende for oss. Det Flanagan her antyder er at dette kun kan skje gjennom å gi et bilde av den moralske aktøren som synes mulig. Å tegne et realistisk bilde av et moralsk ideal kan kun gjøres gjennom å undersøke hvordan ting forholder seg i den faktiske verden. Dermed bør en god moralfilosof kunne gi en redegjørelse for hva som kreves av en aktør for å føre den aktuelle verden nærmere en ideell verden slik den beskrives i en teori. For å gjøre dette på en holdbar måte forplikter teorien seg dermed til prinsippet om minimal psykologisk realisme: "A normative conception which fails to meet certain standards of psychological realizability will fail to grip us, and in failing to grip us will fail to gain our attention, respect, and effort" (Flanagan 1991: 26). Å hevde at en gitt moralteori opererer med en moralpsykologi som ikke er realiserbar i henhold til prinsippet om minimal psykologisk realisme, trenger dermed ikke å føre til at den aktuelle teorien forkastes. Det kan forstås som en ufullstendighet ved teorien fordi en redegjørelse om hvordan prinsipper kan munne ut i handling ikke er utarbeidet. Imidlertid kan det også vise til at en slik redegjørelse faktisk ikke er mulig. I så fall får teorien en utfordring i konkurranse med teorier som kan gi en forklaring på moralsk motivasjon (Flanagan 1991: 38).

Vi har nå sett at en moralteori, for å kunne hevde seg som en seriøs teori, bør kunne gi en redegjørelse for moralsk motivasjon. Vi har også sett at dette gjerne vil gi seg til kjenne i form av en (minimal) psykologisk realisme. Dette innebærer å forplikte seg til å bygge teorien innenfor rammene av hva som synes mulig eller realiserbart for vesener som oss.

Spørsmålet blir nå hvordan man skal avgjøre hva som er psykologisk realiserbart?

Realiserbarhet kan betraktes i to dimensjoner: i forhold til kognitiv struktur og sosialpsykologisk. Disse to dimensjonene kan tolkes som å dreie seg om vår beskaffenhet som naturlige vesener og samfunns-, og kulturvesener. Å skille mellom våre naturlige trekk, og trekk som er et resultat av påvirkning fra samfunn og kultur er notorisk vanskelig, men et skille som vi skal se Flanagan mener er nødvendig å være seg bevisst. Å dele egenskaper inn i sosiale og naturlige trekk er en vanskelig oppgave av flere grunner: Sosiale trekk er betinget og konstruert av naturlige trekk, naturlige trekk er ikke egenskaper som alltid er aktualisert (de er heller å betrakte som muligheter som er avhengig av en viss stimulering for og aktualiseres). Kanskje kan psykologiske egenskaper betraktes som organisert langs et spekter hvor disse er ordnet etter hvor stabile de er i møte med ulike kulturer og samfunn. Egenskaper som viser seg å opptre i relativt uendret tilstand i mange ulike samfunn og til forskjellige tider har vi god grunn til å anta at er naturgitte (men, de utgjør ingen egen klasse). Det vil si, noen psykologiske egenskaper ser ut til å være en del av vår biologiske utrustning. Disse er grunnleggende egenskaper som ikke, eller i alle fall i liten grad, lar seg fjerne eller endre. Flanagan nevner blant annet visse grunnleggende emosjoner, anlegg for sansning, seksualdrift, sult, evne til språk, rasjonalitet og hukommelse som naturlige trekk (Flanagan 1991: 42). For at en moralteori skal være i overensstemmelse med prinsippet om minimal psykologisk realisme må det være et minstekrav at den på en eller annen måte gir rom eller tar høyde for slike naturlige egenskaper. Dette synes å være relativt ukontroversielt. Vi kan være enige om at å kreve at aktører skal radikalt undertrykke eller endre disse egenskapene ville være å stille urealistiske krav. Dette er dermed én dimensjon for realiserbarhet: en moralteori kan være urealiserbar fordi den krever eliminering av egenskaper som tilhører vår biologiske struktur og som ikke kan skrus av eller undertrykkes: ”There are no possible social settings in which persons could live such a life ...” (Flanagan 1991: 46).

Realiserbarhet også opptrer i en annen dimensjon. Vi så at naturlige egenskaper kan være svært vanskelige å forandre eller eliminere. Sosialt variable egenskaper kan også være så godt fundamenterte i oss at de heller ikke lar seg endres eller kompromitteres i særlig stor grad. Flanagan peker på to temaer som synes å være viktig i forhold til dette: De naturlige egenskapene konstituerer og legger grunnlaget for våre sosialt konstruerte egenskaper. Et eksempel er vår naturlige evne til å føle sult. Enhver teori må ta hensyn til at dette er et reelt behov for alle mennesker, og dermed ikke kreve for stor grad av, for eksempel, askese. Som

Flanagan påpeker, er en vanlig strategi for å implementere slike drifter i en teori å lede driften i en bestemt retning eller begrense dem gjennom de handlingene de gir opphav til. Man vil dermed måtte anerkjenne den grunnleggende sultfølelsen, men samtidig kan man innenfor visse kulturer oppfordre til strengt måtehold og opprette normer for hvordan denne driften gir seg til kjenne. Poenget her er at våre medfødte naturlige egenskaper setter visse absolutte grenser for hvordan en moralpsykologi kan se ut. Når disse naturlige egenskapene opptrer i en sosial sammenheng, noe de nødvendigvis må, gir dette opphav til nye, sosialt konstruerte trekk som også kan virke svært begrensende for hva som oppfattes som psykologisk realiserbart. Det oppfattes kanskje ikke som mulig *for meg*, gitt min sosiale og tidsmessige situasjon, å realisere en buddhistisk psykologi hvor askese er sentralt. Det er en psykologi som vil kreve svært mye av meg for å realiseres, kanskje mer enn hva noen vil mene er rimelig at moralen kan kreve. En moralsk personlighetstype kan også være urealiserbar i en enda sterkere forstand, fordi det i en viss sosial og historisk kontekst ikke kan opptre som en reell mulighet. Bernard Williams bruker en parallell til vitenskapshistorien som eksempel: I et valg mellom moderne vitenskapelige teorier om antennelse og tidligere tiders flogiston-teori, vil den sistnevnte ikke være en virkelig mulighet for oss fordi den ikke lar seg forene med hva vi ellers vet om virkeligheten (Williams 1985: 161-162). Gitt en annen bakgrunnskunnskap, slik som den de hadde på 1700-tallet, er det derimot et aktuelt alternativ. Det finnes visse moralske personligheter som er så fremmede for oss at vi ikke kan realisere dem uten å miste grepet om virkeligheten, eller gjøre oss skyldige i massivt selvbedrag (Williams 1985: 160). Ingen av disse personlighetene er noen absolutt umulighet i den forstand at de er naturstridige. Våre egenskaper vil dermed eksistere langs et spekter, hvor noen sosiale trekk vil være vanskeligere å realisere enn andre. Det andre poenget det er verdt å merke seg er at selv om et trekk kan betraktes som naturlig trenger det ikke å være tilstede universelt. Det er avhengig av visse miljømessige faktorer for å opptre som disposisjoner, og disse disposisjonene kan være avhengige av en viss type stimulering for å aktualiseres. Dette munner ut i at vi må anerkjenne at natur og kultur er så vanskelig å skille fra hverandre at et slikt skille kanskje ikke er særlig fruktbart. Dessuten, og kanskje aller viktigst, er det én ting å peke på en egenskap som naturlig, og noe helt annet å hevde at den aktuelle egenskapen har en moralsk verdi. Det ene medfører ikke det andre. En rekke egenskaper, slik som sinne eller egoisme, kan presenteres som naturlige, men det i seg selv vil ikke være nok for å rettfærdiggjøre etiske teorier som konstruerer normer basert på disse egenskapene. Ulike normative teorier er underbestemt av fakta om vår naturlige eller medfødte utrustning. All

oppsamlet kunnskap fra sosiologi, antropologi, psykologi og biologi vil ikke kunne gi oss en "naturlig moral": "... a morality which benefits, or can be read off from, our natural aims, capacities, and interests" (Flanagan 1991: 51). Derimot kan data fra disse fagene gi oss data til et rammeverk for moralen. De vil kunne si oss noe om hva menneskelivet må bestå av, men de gir oss ikke grunnlag for å konstruere en substansiell moral. Gary Watson uttrykker det passende: "An objective account of human nature would imply, perhaps, that a good human life must be social in character. This implication will disqualify the sociopath but not the Hell's Angel" (Watson sitert i Hursthouse 1999: 192). I moralen trenger vi åpenbart mer enn hva vitenskapen kan gi.

2. Psykologisk realisme i praksis

Vi har nå sett at både filosofisk metode og moralteoriens elementer kan (og bør) la seg påvirke av funn i empiriske vitenskaper. Vi så hvordan en slik påvirkning allerede er et faktum fordi moralen dypest sett krever at vi tar empiri med i beregningen. I tillegg kan det betraktes som noe å strebe etter: Først, rent metodisk fordi filosofien ikke har en egen distinkt a priori metode. For å unngå feilaktige slutninger og filosofiske illusjoner bør empiriske hensyn være noe man retter seg etter. Deretter fordi psykologiske strukturer som er empirisk adekvate i større grad unnslipper problemer med å sikre aktørens motivasjon. De opererer med et normativt *bør* som vi kan forstå og kjenne oss igjen i fordi vi kan relatere den til vårt levesett og problemer vi møter i våre liv. Det vil si, en moral som er "realizable by human creatures like ourselves, and applicable to the kinds of problematic situations we encounter in our lives" (Johnson 1996: 49). Fremfor å ta stilling til forholdet mellom moralske dommer og motivasjon som eksternalisme eller internalisme vil en naturalist se på hva det er som motiverer oss, og hvordan motivasjon faktisk finner sted. Menneskelig motivasjon lar seg kanskje ikke fint innordne etter humane kategorier. En naturalist vil likevel betrakte de faktiske psykologiske mekanismene som ligger bak våre dommer som rettesnorer når vi konstruerer teorier.

I hvor stor grad empiri skal vektlegges er et viktig spørsmål. Filosofien finner noe av sin egenart i å ha en viss kritisk distanse til omgivelsene. Vi så dette hos Prinz, og også Flanagan fremhever at filosofi innebærer en form for refleksjon som gjøres med en viss avstand til omverdenen: "... philosophical psychology has certain traditional advantages in the realm of breadth, generality and systematicity" (Flanagan 1991: 16). Hvor generaliserende kan en filosof tillate seg å være? Hvilket abstraksjonsnivå er det passende å befinne seg på? I stor grad tror jeg dette er spørsmål som ikke lar seg avgjøre uavhengig av praksis. Noen filosofiske disipliner, slik som metafysikk, tillater at man er svært abstrakte og at man opererer relativt fjernt fra betraktninger om hvordan vi erfarer verden⁹. Etikk og

⁹ Enkelte deler av fysikken håndterer størrelser som ikke lar seg testes empirisk, og konstruerer svært abstrakte teorier (strengteorien er et eksempel). Alt dette foregår uten at fysikken mister sin status som empirisk vitenskap. Poenget mitt her er at innenfor de fleste vitenskaper vil jeg anta at vi kan finne et slikt spenn mellom det som er spesifikt, konkret og empirisk, og noe som er mer generaliserende og abstrakt. Også er her filosofi på lik linje med andre vitenskaper.

moral derimot er emner som det er nærliggende å tro at krever kunnskap om hva et menneske er, og dermed større bevissthet rundt empiriske spørsmål. Fordi en passende kalibrering mellom det abstrakte og det konkrete kanskje ikke lar seg avgjøre på prinsipielt grunnlag vil jeg kaste lys på dette temaet ved hjelp av en pågående debatt innen moralfilosofien. Denne debatten omhandler hvorvidt dydsetikken er empirisk holdbar. Mer spesifikt dreier den seg om vi kan hevde at dydene finnes som psykologiske strukturer. Nyere sosialpsykologisk forskning gir oss grunn til å tro at dyder ikke eksisterer i en form som gir dydsetikeren tilstrekkelig grunnlag for sine påstander.

Det er flere grunner til at jeg trekker frem akkurat denne debatten. Først, dydsetikken fremstår som en teori om moral som er spesielt åpen for empirisk påvirkning. Dydsetikerne liker å påpeke at deres moralpsykologi er mer realistisk enn den som fremmes av filosofer som baserer seg på plikt og prinsipper. Det er dessuten vanskelig å forstå dydene på noen annen måte enn som psykologiske strukturer som på en eller annen måte lar seg realisere. Dydene synes å være begreper som balanserer mellom det konkrete og empiriske, og det mer abstrakte og generelle:

... moral psychology need not be "merely psychological" – talk of virtue is not usually supposed to be reducible to purely psychological terms. But the virtues would not be able to do their job if they did not possess or engage an interconnected set of causal powers. These powers would help explain why the virtuous act, or see things, as they do. (Railton 1995: 94).

For at dydene skal kunne fylle sin sentrale funksjon må de ta del i alminnelige kausale kjeder, og dessuten selv ha kausal makt. De er ment å utgjøre et grunnlag for å anslå hvordan aktører vil handle: vi kan forvente at en ærlig aktør forteller sannheten. Dydene er også ment å forklare aktørers handlinger: den ærlige forteller sannheten fordi han verdsetter sannhet. Avstanden mellom personlighetspsykologi og dydsetikk fremstår ikke som veldig stor. Det er dermed heller ikke umiddelbart urimelig å måle dydsetikken opp mot de samme standardene vi forventer at personlighetspsykologien lever opp til.

En annen grunn til å ta tak i denne debatten er at den viser hvordan man kan være empirisk filosof i praksis. John Doris, som fungerer som talperson for kritikerne av dydsetikken i denne oppgaven, er en empirisk filosof slik Prinz beskriver dem: Han henter inn data fra ulike forskningsfelter innenfor psykologien og bruker disse til å kritisere sentrale filosofiske

forståelser av bevissthetens oppbygning og hvordan den fungerer. Doris baserer seg i stor grad på sosialpsykologisk forskning, og på grunnlag av disse dataene betviler han eksistensen av visse personlighetsstrukturer som forekommer i filosofisk teori, nemlig dydene. Han er tilstrekkelig abstrakt og generell (han har satt seg godt til rette i lenestolen) til å bedrive filosofi, men lar sin tenkning lede av empirisk forskning. Slik oppsummerer Prinz virksomheten til empiriske filosofer: "Empirical philosophers tend to be interested in relatively general, first-order questions, and, when they analyze concepts, they typically do so from the armchair in order to guide and systematize empirically informed investigations of how the mind works" (Prinz 2008: 199). Når Doris kommer med egne forslag om hvordan man skal forstå vår moralske praksis gjør han dette ved å vise hvordan empirien, gitt en viss tolkning, kan stemme overens med innholdet i sentrale moralske begrep eller at moralens begreper lar seg tolke på en måte som er kompatibel med empiriske funn. Som empirisk filosof bedriver Doris "... armchair conceptual analysis to guide or interpret first-order empirical results" (Prinz 2008: 198). Debatten er dermed ikke kun av interesse fordi den viser hvordan tradisjonelle filosofiske begrep kan være uholdbare i lys av empiri, men også fordi den belyser en ny trend hvor man åpner opp filosofien for andre forskningsfelter. Prinz kaller det *metodisk pluralisme*.

Owen Flanagan var en av de første som brakte på banen en empirisk basert kritikk av dydsetikken. Helt konkret går kritikken ut på at en rekke studier har vist at vi ikke er konsistente og forutsigbare nok i våre handlingsmønstre til å kunne tilskrives karaktertrekk¹⁰. Handlingene våre produseres under stor grad av påvirkning fra situasjonene vi befinner oss i, fremfor å være produkter av indre strukturer alene. Dette gjør at Flanagan mistenker at noen felter av moralpsykologen ikke tillater store generaliseringer eller forenklinger:

Moral psychology may be less unified than we typically think ... because our moral dispositions and abilities are of many different types, with different learning histories, different relations to temperament and rationality, and different susceptibilities to different kinds of external forces. (Flanagan 1991: 286)

¹⁰ Det vil si karaktertrekk i en tradisjonell forstand: "... by 'character' we mean a settled way of being, which is largely internal to the agent and which is relatively immune to situational disturbances." (Flanagan 1991: 264).

Kanskje kan vi ikke vurdere handlinger og aktører uavhengig av situasjoner? Pluralisme er altså ikke bare et relevant stikkord i forhold til filosofisk metode, men også i forhold til hvordan man skal betrakte strukturelle aspekter ved bevisstheten. Hvis dydsetikerne betrakter dydene som relativt isolerte fra omgivelsene, vil de få problemer i lys av hva Flanagan her påpeker. Striden kjerne er altså om dydsetikken er sårbar for empirisk kritikk, og hvorvidt de kan ta disse dataene alvorlig og samtidig beholde teoriens sentrale elementer. For å forstå debatten må vi se på hvordan dydsetikken beskriver dydene, og hvordan de selv streber etter å være psykologisk realistiske. Dette vil være tema gjennom de neste sidene.

2.1 Dydsetikk og empiri

Dydsetikken fremmes gjerne som en teori som har en mer plausibel moralpsykologi, i form av en bedre redegjørelse for moralsk motivasjon, enn sine konkurrenter. Rosalind Hursthouse er kritisk til det dogmatiske skillet mellom ønsker og trosoppfatninger som står sentralt i de fleste moderne teorier om (moralisk) aktørskap. Hun fremlegger overbevisningen om at disse elementene virkelig er bestanddelene i aktørers moralske valg og handlinger som en uutalt antagelse alle deler. Ønsker og trosoppfatninger regnes som absolutt adskilte, og som *natural kinds*, altså som størrelser som faktisk eksisterer (Hursthouse 1999: 16). Denne måten å betrakte menneskelig psykologi har gitt filosofer mye hodebry. Problemet oppstår fordi Humes modell synes å gjøre visse sentrale aspekter ved vår moralske praksis inkompatible. Problemets kjerne er hvordan vi skal forklare at moralske dommer motiverer, og samtidig hevde at moralske dommer er objektive. Kun ønsker er motiverende for oss, og disse lar seg angivelig ikke påvirke av trosoppfatninger. Ikke overraskende argumenterer Hursthouse mot Humes modell. Heller enn å operere med absolutte skiller når det gjelder moralsk motivasjon burde vi tenke motivasjon som noe som fremtrer i ulike grader.

Hursthouse er på ingen måte noen pioner i å utfordre moderne moralteorier: Dydsetikere har kritisert deontologer og konsekvensialister for deres moralpsykologi i over et halvt århundre. I ”Modern Moral Philosophy” (1958) anklager Elizabeth Anscombe sine kolleger for å benytte seg av en moralpsykologi som er uholdbar: ”In present-day philosophy an explanation is required how an unjust man is a bad man, or an unjust action is a bad one; to give such an explanation belongs to ethics; but it cannot even be begun until we are

equipped with a sound philosophy of psychology” (Anscombe 1958: 4). For å frigjøre moralfilosofien fra problemer slik som Moores ”Open-Question Argument”, debatten om internalisme og eksternalisme, og teorier som åpner opp for kontraintuitive moralske dommer mener Anscombe at dydene, slik vi finner dem hos Aristoteles, må bringes tilbake på banen. Kun gjennom å presentere rettferdighet som dyd kan man gi en tilfredsstillende redegjørelse for moralske begreper av denne typen. Først må man klargjøre hva slags egenskap en dyd er og forholdet mellom dyd og handling: ”For this we certainly need an account at least of what a human action is at all, and how its description as ’doing such-and-such’ is affected by its motive and by the intention and intentions in it ...” (Anscombe 1958: 5). Forutfor teoribygging trenger vi et fundament bestående av en moralpsykologi som holder mål. Anscombe redegjør ikke for hvorfor hun mener moralpsykologi må utarbeides forutfor teori, og hun gir heller ingen forklaring på hvordan en moralpsykologi som baserer seg på dyder må være. Likevel, det som ble brakt frem i hennes artikkel ble i stor grad tatt opp i samtidens moralske diskurs, og det mangler ikke på aktører som i ettertid har redegjort for dyd i mer detalj. Dydsetikken ble brakt tilbake i rampelyset nettopp fordi den ble ment å fremme en mer holdbar moralpsykologi.

Vi kan se tilbake på tema fra forrige kapittel, nemlig psykologisk realisme. Dydsetikerne mener både at de standardene som gis av PMPR er verdt å strebe etter, at konkurrerende moralteorier mislykkes i en slik streben, og at de selv gir en redegjørelse for aktør og moral som innebærer en høy grad av psykologisk realisme. Sitatet fra Anscombe ovenfor inviterer nærmest vitenskapen inn i debatten. På en side lykkes de med å fremstå som psykologisk realistiske og de kan også komme med empirisk belegg for deler av sin posisjon¹¹: Å betrakte aktøren i sin helhet fremfor kun å fokusere på enkelthandlinger gjør at vi kan bevege oss bort fra et ensidig fokus på rasjonalitet, domfelling og moralske prinsipper til et mer åpent perspektiv som også innlemmer følelser, affekter og reell motivasjon (Railton 2009: 10). Studier gjort innenfor kognitiv psykologi støtter faktisk opp under et slikt syn. I nyere forståelser av hvordan vi organiserer kunnskap er kunnskap tenkt konstruert rundt en prototype-basert modell. I stedet for å anta at kunnskap om begreper og begrepsanvendelse er organisert som språklige definisjoner med nødvendige og tilstrekkelige betingelser for anvendelse (den klassiske modellen), antar man at kunnskap er organisert som en base med

¹¹ Jeg kjenner ikke til hvorvidt noen dydsetikere faktisk har forsøkt å komme med slikt empirisk belegg. Forskningen som det her refereres til er det jeg som trekker frem som relevant.

oppsamlede modeller som fungerer som prototyper ved begrepsanvendelse. En prototype er en konstruert enhet som samler opp de statistisk sett vanligste kjennetegnene ved konkrete gjenstander vi eksponeres for. Fysiske gjenstander trenger ikke å ha alle kjennetegnene vi finner i prototypen for å falle inn under et begrep. Forskning i kognitiv vitenskap viser at vi heller antar at begreper stemmer overens med konkrete gjenstander eller hendelser *i større eller mindre grad* (May, Friedman og Clark 1996: 5). Moralsk kunnskap og moralske begreper skiller seg ikke fra andre begreper i denne modellen. Konsekvensen av dette er at det fremstår som relativt lite plausibelt å mene at man kan oppsummere eller kondensere moralsk kunnskap i prinsipper eller moralske regler. Årsaken til dette er at vårt oppsamlede lager av prototyper vil, hos et voksent menneske, være svært stort. *For* stort til å på en tilfredsstillende måte la seg oppsummere i form av enkle regler. Her følger dydsetikken vitenskapen i å argumentere mot en forståelse av etikk som regel eller prinsipp-basert.

To hope to condense the subtle moral expertise encoded by a biologically realistic neural network into a tractable set of summary principles or moral rules is in all probability a quite hopeless task. In this respect, moral expertise looks set to follow the pattern of expertise in general. Long volumes fail dismally to capture the knowledge of the expert chess player. The knowledge of the successful moral agent is probably no less complex, and no more amenable to brief linguaform summation (May, Friedman og Clark 1996: 7).

Interessant er det også at det her brukes analogier mellom moralsk praksis og utøvelsen av en ferdighet, en analogi som er svært sentral for dydsetikerne. Dessuten er moderne psykologiske teorier om hvordan resonering og overveielse foregår konstruert slik at det ikke er noen radikale forskjeller mellom moralsk overveielse og andre typer tenkning. Begge deler kan beskrives som en evne til å gjenkjenne mønstre og generalisere over disse mønstrene (for eksempel, å ordne gjenstandene rundt oss under bestemte prototyper, og å konstruere disse prototypene). Grunnleggende evner, slik som sansning, og høyerestående evner, slik som evnen til å gjøre moralske vurderinger, er organisert langs et kontinuum: de er av samme type (May, Friedman og Clark 1996: 7). Moralsk ekspertise kan betraktes som en særlig form for perseptuell skarpsynthet:

Moral failure and moral insight are thus explained as, in effect, rather subtle cases of perceptual failure and perceptual acuity. The good agent commands a set of developed

moral prototypes that structure the way she processes inputs and hence, in a profound sense, the way she sees the world (May, Friedman og Clark 1996: 7).

En slik måte å betrakte moralsk klokskap på vil også kunne produsere en viss gjenkjennelse hos dydsetikeren. Dydsetikken kan dermed med en viss grad av selvtillitt hevde å være en filosofisk teori om moral som også innlemmer viktige sider ved menneskelig psykologi. Lykken er imidlertid kortvarig, en del av dydsetikken har i den senere tid blitt kritisert for å benytte seg av en psykologi som er urealistisk, nemlig den delen som omhandler dydene som robuste karaktertrekk som bør kultiveres.

2.2 Dydene som karaktertrekk

Et naturlig sted å begynne er å se nærmere på hva dydene er. Både Julia Annas og Rosalind Hursthouse, kjente representanter for en neo-aristotelisk dydsetikk, mener klart at dydene skal forstås som karaktertrekk. Annas skriver hvordan Aristoteles mente dydene var *hexis*, en tilstand i menneskets sjel. Annas mener videre at disposisjon er et bedre moderne ord for Aristoteles' *hexis* enn tilstand: ” ... a virtue like courage is a disposition because it is a condition because of which I am so disposed as to act in brave ways; and this is what a *hexis* is” (Annas 1993: 50). Dydene er altså disposisjoner til å handle på bestemte måter (eller enda sterkere: de er betingelser for å handle på disse måtene). Annas tillegger også dydene et affektivt og intellektuelt aspekt, noe som innebærer at de er noe mer enn rene vaner. Dydene presenteres som en type karaktertrekk, et kompleks av emosjoner, kognisjoner og adferdsdisposisjoner satt sammen på en bestemt måte,

... having a virtue is having one's character developed in such a way that one not only grasps what the right thing to do is but takes pleasure in doing it; one is repelled by the thought of wrong acting; and one is not seriously tempted by incentives not to do the right action. Virtue is a state of the agent's character and emotions, not merely a disposition to act in certain ways. (Annas 1993: 55)

Her ser vi hvordan det vektlegges at dydene både har et ytre og et indre aspekt. Det er slikt sett en pakkeløsning, hvor både handling og bestemte sinnstilstander er nødvendig. Å handle dydig er å handle ut fra en bestemt type grunner. Også hos Hursthouse finner vi dydene presentert som karaktertrekk: ” ... a virtue is generally held to be a character trait, a state of

one's character" (Hursthouse 1999: 11). På samme måte som Annas vektlegger hun både den ytre handlingen og de indre prosessene. Hursthouse poengterer videre at nettopp fordi dyd er noe mer enn kun handlingsdisposisjoner er de relativt lite utsatt for forandring. Dyd er noe som involverer hele vår kognitive struktur, og innebærer forutsigbar adferd av en bestemt type. Vi forventer konsistens i den adferden som gjenspeiler de ulike dydene (Hursthouse 1999: 12). Dessuten fremhever Annas at dydene må betraktes med henblikk på både fortid og fremtid. Dydene er stabile disposisjoner, og er resultater av gjentatte beslutninger og handlinger i aktørens fortid, og vil gi seg til kjenne i fremtidig adferd. En gitt handling avslører både hvem vi er, og har betydning for hvem vi vil bli: "We are all the time faced with new and complex situations; how we deal with them reveals what we have become and affects what we are becoming" (Annas 1993: 52). Hvordan vi handler i ukjente situasjoner og hvordan vi handler over tid har dermed stor innvirkning for hvorvidt man er dydig eller ikke. Å være dydig er ikke en egenskap man kan være i besittelse av i en situasjon, men mangle i en annen.

Anscombe fremmet dydsetikken fordi man på denne måten fanget opp hele aktøren som moralsk relevant, i motsetning til andre teorier som kun omfattet enkelthandlinger. Dydsetikere liker å sammenligne tilegnelsen av dydene med hvordan man lærer seg praktiske ferdigheter. Dydene er noe man kommer i besittelse av gjennom læring under påvirkning av en eller annen form for rollemodell. Gjennom å tilegne seg dydene vil man i økende grad bli selvstendig og i stand til å reflektere over moralske emner. Det er en prosess hvor man går fra å være "elev" som følger regler, til å være utlært og kunne komme med mer generelle betraktninger og samtidig ta hensyn til det partikulære i ulike situasjoner. Man gjenkjenner og er sensitiv ovenfor sentrale perspektiver i enkeltsituasjoner, og dette gjøres på grunnlag av en dypereleggende og mer vidtrekkende forståelse av hva som kreves av den dydige. Å være dydig innebærer dermed å være i besittelse av en bestemt type kunnskap, og en spesiell strukturering av denne kunnskapen. Først kan vi merke oss at for å handle dydig er det en forutsetning at de rette grunnene, altså den rette kunnskapen, er tilstede i aktøren. Man kan ikke være både ond og dydig eller være dydig ved en tilfældighet, "... there is an important moral difference between the person who merely acts rightly and the person who is wholehearted in what she does" (Annas 2006: 517). Moralsk verdi for dydsetikeren avgjøres altså av noe mer enn kun den observerbare handlingen. Aktørens indre strukturer spiller en viktig rolle. Dette bringer meg videre til et annet viktig punkt, nemlig at moral for dydsetikeren ikke består av enkeltstående moralske påbud i form av et sett med regler.

Moralen forstås som en bestemt type kunnskap, eller slik Annas uttrykker det: kunnskap som forståelse. Denne type kunnskap kan kontrasteres mot kunnskap tenkt som en forbedret versjon av fakta man før bare hadde en antagelse om (Annas 2001: 242). Dette er et tema vi kjenner igjen fra Platons dialoger: kunnskap handler om å kunne gjenkjenne det generelle i det partikulære, og på denne måten være i stand til å ordne enkelterfaringer under større og mer enhetlige kategorier. Moralsk kunnskap består av en forståelse som bygges opp gjennom erfaring og opplæring. Også Anscombe minner oss på at moral ikke lar seg bokstaveres ut i enkle prinsipper eller beskrivelser vi kan klistre på enkelthandlinger: "We should no longer ask whether doing something was 'wrong', passing directly from some description of an action to this notion; we should ask whether, e.g., it was unjust; and the answer would sometimes be clear at once" (Anscombe 1958: 9). Fremfor å basere seg på regler, baserer dydsetikken seg på forståelse av moralens fordringer, en forståelse som bygges opp over tid. Den dydige handlinger gjenspeiler ikke bare enkeltstående valg, men et helt sett med karaktertrekk som bygges og opprettholdes gjennom de valgene man har tatt og tar gjennom livet. Dydsetikerne har dermed en ganske annerledes forståelse av rasjonalitet enn den som ofte gjør seg gjeldende i moralfilosofien i det tjuende århundret, nemlig den instrumentelle. På grunn av dette perspektivet mener dydsetikere å kunne gi en mer realistisk redegjørelse for moralsk motivasjon. Dydene består av en bestemt type resonering og valg gjort på grunnlag av en bestemt type grunner. Denne resoneringen og disse valgene mener Annas tar sitt utgangspunkt i en type tenkning vi alle er i besittelse eller potensielt i besittelse av: "... virtue ethics tries to improve the reasoning we all share, rather than replacing it by a different kind" (Annas 2006: 516). De påberoper seg altså å gi et mer psykologisk realistisk bilde av hvordan moralen opererer i oss, og at det er allerede eksisterende psykologiske strukturer som utgjør potensialet for forbedring.

Vi kan også koble dydsetikken opp mot forrige kapittels redegjørelse for psykologisk realisme gjennom å se nærmere på hvordan dydsetikeren mener å kunne fundamentere moralen i menneskelivet på en bedre måte enn konkurrerende teorier. Dydene er ikke personlighetsstrukturer som utelukkende har en moralsk funksjon. Dydene omfatter hvem vi er i en vid forstand, vår identitet, våre verdier og forholdet til omgivelsene våre. Eller, kanskje mer korrekt, gjennom dydsetikken får disse aspektene også en plass i moralteorien. Vi så hvordan Flanagan mente at en psykologisk realistisk teori må innlemme slike aspekter som ikke har noen direkte moralsk relevans. Gjennom dydene kan moralen kobles opp mot andre elementer som er viktige for oss, slik som å leve liv som oppfattes som verdige og av

betydning (for oss og andre): kort og godt, liv som er verdt å leve. Dette, påpeker Hursthouse, er elementer som ikke gis rom i andre moralteorier. Deontologien og utilitarismen "...ignored or sidelined a number of topics that any adequate moral philosophy should address". Blant elementer som mangler i disse to teoriene finner Hursthouse (i tillegg til en akseptbar redegjørelse for moralsk motivasjon): "...moral education, moral wisdom or discernment, friendship and family relationships, a deep concept of happiness, the role of the emotions in our moral life, and the question of what sort of person I should be, and of how we should live" (Hursthouse 1999: 3). Dette er elementer som gir moralen rotfeste i menneskelivet. Moralene oppfattes som en del av meg, i motsetning til noe fremmed eller tvunget jeg må rette meg etter.

Denne gjennomgangen viser altså at dydene er å betrakte som karaktertrekk: De er stabile, indre strukturer som er kausalt effektive. Dydsetikken er slik sett ikke kun en teori som kommer med isolerte påstander om moral. Det fremmes substansielle påstander om hvordan våre handlinger skal redegjøres for, hvordan vi tilegner oss visse typer kunnskap, hvordan denne kunnskapen lagres i oss, og at kultivering av visse karaktertrekk er et sentralt element for at vi skal trives. Som vi så i kapittel 2.1 har noen av disse påstandene godt empirisk hold. Tanken om kunnskap som prototypebasert kjenner vi igjen her¹².

Dydsetikken er altså tilstrekkelig deskriptiv til at sentrale elementer av teorien lar seg empirisk teste. Dydsetikken tar også sitt utgangspunkt i hvordan vi forstår oss selv, i intuisjoner vi har om hvem vi er, og hvordan tar avgjørelser og handler. Å betrakte og beskrive seg selv og andre ved hjelp av karaktertrekk synes å ha en bred intuitiv apell (for oss vestlige). Problemet er at vi ikke alltid kjenner oss selv best. Sosialpsykologiske studier problematiserer karakterbegrepet, og vi skal se at dette skaper problemer for dydsetikkens mulighet til å hevde seg som en psykologisk realistisk teori.

2.3 John Doris' kritikk av dydsetikken

Vi har til nå sett på hvordan dydsetikerne fremmer sin teori blant annet på grunnlag av å kunne tilby en spesielt plausibel moralpsykologi. Noe av grunnen til dette er at de mener å

¹² Det må presiseres at dydsetikerne ikke selv har bygd sin teori i tråd med slik forskning. Samsvar her er mer å regne som en lykkelig tilfældighet.

grunnlegge sin teori i menneskelivet slik det faktisk leves, og å kultivere en tenkemåte vi allerede i større eller mindre grad er i besittelse av. Dydsetikken oppstår som en teoretisering over refleksjoner vi alle gjør, påstås det (Annas 2006: 526). La oss anta at i vår hverdagslige omgang med andre mennesker tenker moral blant annet gjennom karaktertrekk. At dydsetikken på denne måten korresponderer med våre intuisjoner om vår moralske og psykologiske virkelighet. Vi tillegger aktører (og oss selv) karaktertrekk som er ment å beskrive hvordan vi handler, og også si noe om hvilken grad vi lever opp til moralens fordringer. Hva hvis dette bildet er feil? Hva vil det innebære hvis det viser seg at denne måten å reflektere over moralske spørsmål ikke stemmer overens med hvordan vi faktisk opererer som aktører i virkeligheten? Mer konkret: Hvilken betydning har det for dydsetikken hvis dydene, altså karaktertrekkene, som er teoriens mest sentrale element, ikke kan regnes som reelle størrelser? At de ikke korresponderer med slik vi faktisk handler? At det ofte er på urettmessig grunnlag at vi tilskriver karaktertrekk til andre aktører? Dette er spørsmål som dydsetikken har blitt utfordret med i kjølvannet av populariteten den oppnådde i siste halvdel av 1900-tallet. Det er blitt påpekt at å postulere karaktertrekk slik dydsetikerne gjør er problematisk fordi slike trekk ikke har latt seg påvise med utgangspunkt i faktisk adferd: systematisk observasjon viser at menneskelig handling ikke er strukturert på den måten de som fremmer en moralpsykologi basert på karaktertrekk påstår (Doris 2005: 632). Utfordringen ligger dermed i at dydsetikkens sentrale elementer er empirisk uholdbare.

John Doris er antageligvis den som har bokstavert ut denne tematikken mest fullstendig. Hans bok, *Lack of Character*, fra 2002 vil være mitt utgangspunkt når vi ser nærmere på debatten. Doris' poeng er at dydsetikken kommer med visse påstander om aktørenes psykologi og adferd. Disse påstandene lar seg undersøke empirisk, og det er, mener Doris, i utgangspunktet ingen grunn til å tro at dydsetikeren ikke skulle åpne for dette. At karaktertrekk, her: dyder, gir opphav til konsistens og forutsigbarhet i en aktørs adferd er et sentralt punkt for dydsetikerne. Vi kan med trygghet ha visse forventninger til den dydige: "Most obviously we expect a reliability in their actions" (Hursthouse 1999: 10). Dette plukker Doris opp: "Attributions of character and personality traits is associated with behavioral expectations" (Doris 2002: 16). Dessuten, mener Doris, er man innenfor dydsetikken også opptatt av å komme med forklaringer på aktørenes handlinger. De ønsker å si noe om hvilke motiver, følelser og tanker som ligger til grunn for visse handlinger. Å påstå at noen gjorde noe modig er å henvise til en psykologisk tilstand i aktøren som konstitueres av bestemte typer psykologiske elementer. Dydsetikken lar seg lese ikke bare

evaluativt: Å forstå dydsetikerens påstander som deskriptive vil ikke være å misforstå dem. Innenfor eksperimentell sosialpsykologi det siste århundret har det blitt gjort en rekke studier av karaktertrekk. Disse studiene synes å ha en direkte relevans for hvorvidt dydsetikerene kan betrakte dydene som psykologisk realistiske.

Vi så tidligere hvordan dydene skal forstås som karaktertrekk. Innenfor sosialpsykologien har det i det vokst frem en tradisjon som på grunnlag av en rekke empiriske studier stiller seg kritisk til visse måter å betrakte karaktertrekk og personlighet på. Det disse studiene utfordrer er den alminnelige tanken om at det meste av hva en aktør foretar seg kan forklares ut fra personlighet og tidligere adferd, samt at karakter og personlighet er godt grunnlag for å forutse hvordan en aktør vil handle i en gitt situasjon. Innenfor folkepsykologien (og moralpsykologien som mange filosofer opererer med) er det å ty til personlighet, disposisjoner, karaktertrekk og, i visse tilfeller, dyder den vanlige måten å strukturere, og gi mening til våre egne og medmenneskers handlinger på. Forskere, derimot, kan peke på at gjentatte studier viser oss at denne måten å forklare handling på er en feilslått strategi. For det første: I de fleste tilfeller vil forutsigelser gjort på grunnlag av en aktørs disposisjoner ikke være vellykket. Karaktertrekk synes å være svært lite kausalt effektive (i alle fall når det gjelder overgangen fra tanke til handling), spesielt når aktøren befinner seg i nye og ukjente situasjoner. Hvis koblingen mellom handling og karaktertrekk mangler vil de fleste være tilbakeholdne med å tildele en aktør et karaktertrekk. Det er en beviselig stor forskjell mellom den konsistensen i adferd som vi må anta at forutsettes av folkepsykologiens og dydsetikkens forståelse av karakter, og den konsistensen som faktisk kan observeres (Doris 2002: 20). Generelt sett kan man påstå at ingen av oss utviser tilstrekkelig konsistens i adferd til å kunne legitimere tilskrivningen av et karaktertrekk. Tendensen til å trekke slike ugyldige slutninger på grunnlag av karaktertrekk har fått navnet *the fundamental attribution error*: "People's inflated belief in the importance of personality traits and dispositions, together with their failure to recognize the importance of situational factors in affecting behavior..." (Ross og Nisbett 1991: 4). Doris mener likevel at vi med rettmessighet kan gjøre noen generaliseringer, men i mye mindre grad enn hva vi vanligvis tillater oss når vi tenker på karaktertrekk. Blant annet må vi være forsiktige med, og kanskje til og med i de fleste alminnelige tilfeller avstå fra, å gjøre generaliseringer på tvers av situasjoner. I forhold til vår bruk av karaktertrekk i sin alminnelighet mener Doris vi kan snakke om *overattribuering*: Vi har en tendens til å gjøre slutninger fra handlinger til besittelse av karaktertrekk på et for tynt grunnlag (Doris 2002: 101). Denne tendensen innebærer også at

vi undervurderer et element som studier har vist at har stor innvirkning på adferd, nemlig situasjonen man befinner seg i. I de fleste situasjoner hvor vi gir karaktertrekk "æren" for handlinger burde vi heller kreditere omgivelsene. Situasjonisme er navnet man har gitt den tradisjonen som på grunnlag empirisk forskning fremhever at omgivelsene, fremfor indre strukturer, har størst betydning for hvordan vi handler.

A half century of research has taught us that in ... novel situations, one cannot predict with any accuracy how particular people will respond. At least one cannot do so using information about an individual's personal dispositions or even about that individual's past behavior. (Ross og Nisbett 1991: 2)

Det som gir konsistens i adferd er ikke karaktertrekk, men elementer man gjenfinder i situasjonene man handler i. Litt forenklet kan man kanskje påstå at hvorvidt vi er ærlige eller ikke, ikke avgjøres av en disposisjon vi har (eller mangler), men av situasjonelle elementer som virker fremmede for denne typen adferd: "... honesty is not an 'inner entity' but is instead 'a function of the situation'" (Doris 2002: 24).

Doris påstår ikke at alle måter å tenke karaktertrekk på er empirisk uholdbare, og retter sin kritikk mot én bestemt betraktning av karakter. Doris ønsker å vise at de som betrakter karakter som et sett med robuste trekk som er evaluativt integrerte¹³ mangler empirisk grunnlag for sin posisjon. *Globalisme* er navnet har gir denne posisjonen, det er en slik forståelse av karaktertrekk vi finner i dydsetikken. Julia Annas fremlegger globalismen på følgende måte: "When, that is, we ascribe a virtue, we are ascribing a disposition which is robust, producing brave actions in a reliable way and one which is relevant to Jane's life overall, and thus global" (Annas 2003: 2). Et robust karaktertrekk vil kunne tilskrives en aktør hvis han kan utvise *konsistens*, altså at det aktuelle trekket kommer til kjenne i en rekke ulike situasjoner, og *stabilitet*, noe som innebærer at trekket gir utslag i distinkt adferd over tid (Doris 2002: 22-23). Ifølge Doris har vi godt empirisk grunnlag for å hevde at konsistente karaktertrekk ikke finnes. Derfor mener Doris at globalisme ikke er en holdbar måte å betrakte vår personlighet og psykologiske struktur på.

¹³ Evaluativ integrasjon er hva vi i dydsetikkens termer kjenner som dydenes enhet. Det vil si, at du ikke kan ha én dyd uten å ha dem alle. Doris legger frem bevis for at også dette er en empirisk uholdbar påstand. Jeg oppfatter kritikken av dyder som karaktertrekk som mer fundamental: finnes det ikke dyder kan det heller ikke oppstå noen enhet mellom dem. Derfor lar jeg denne innvendingen fra Doris ligge.

2.4 Annas' respons

Det har naturlig nok oppstått en debatt om dyder og empirisk holdbarhet i kjølvannet av Doris' utspill. Mitt håp er at denne debatten vil belyse sentrale temaer som oppstår i møtet mellom empiriske vitenskaper og filosofiske teorier. Jeg kan umulig komme med en komplett redegjørelse som favner alle sider og innspill i debatten, men jeg har valgt ut noen artikler jeg tror er representative. Jeg bruker i første rekke en serie artikler fra *Philosophy and Phenomenological Research* som tar utgangspunkt i *Lack of Character*. Her får Doris selv svare på kritiske innvendinger fra Julia Annas, Nomy Arpaly og Robert C. Solomon.

Annas er den som først entrer scenen for å kommentere Doris' bok. Annas er en av de dydsetikerne som har vært mest profilert i denne debatten, og for å yte henne rettferdighet vil jeg ikke kun bruke teksten fra det aktuelle temanummeret, men også en annen tekst hvor hun har en lengre redegjørelse for sin posisjon (se Annas, 2003). Dessuten vil jeg trekke inn en artikkel av Gopal Sreenivasan som tar opp en innvending av samme type som Annas'.

Annas innvending er spennende å se nærmere på fordi den omhandler i hvilken grad empirisk forskning er relevant for, eller på en tilfredsstillende måte kan brukes til å vurdere, visse filosofiske begreper. Annas stiller seg spørrende til hvorvidt sosialpsykologiens eksperimenter er konstruert på en slik måte at de egner seg for "loggføring" av dydig handling. Eksperimentene Doris bruker begrenser seg til observasjon av adferd, og utelater et perspektiv som er viktig for å kunne påstå at noen er i besittelse av dyd, nemlig grunnene som ligger bakenfor handlingene. Disse forsøkene gir ikke rom for noen forståelse av den rollen praktisk rasjonalitet spilte hos deltagerne:

An observer, particularly a scientist setting up an experiment, is recording behavior of a certain kind: picking up money, copying in a test and so on. And there are many ways in which this kind of observation may fail to track what is important for virtue, namely the kind of reasons acted on. (Annas 2003: 13)

Er dydene i det hele tatt noe som lar seg empirisk studere kun fra en observatørs perspektiv? Dette er et tema jeg vil ta opp. Først vil jeg legge veien om en annen og relatert innvending, nemlig om Doris kanskje kan anklages for å tegne et feilaktig bilde av dydene nettopp fordi han lener seg på denne typen forskning.

Både i valg av empiri og i hvordan Doris redegjør for dydene mener Annas at han mislykkes i å gi en redegjørelse som fanger opp de indre strukturene som konstituerer dyd. Ifølge Annas tegner Doris et bilde av den dydige som en aktør som har tilegnet seg de dydige trekkene i form av ukritiske og fastlåste vaner. En slik (feil)konstruksjon av dydene gjør at vi forventer at den påstått dydige vil kaste seg ut i enhver situasjon for å gjennomføre en på forhånd utarbeidet handlingsplan (Annas 2005: 637)¹⁴. Dyd slik forstått handler mer om ren viljestyrke og handlekraft enn utøvelse av praktisk rasjonalitet. Ifølge Annas utelater Doris hvordan den dydige alltid tar hensyn til det partikulære i hver enkelt situasjon, og hvordan den dydige er i stand til å oppfatte hva som er det rette å gjøre, ikke fordi hun følger en forutbestemt plan, men fordi det er hva den aktuelle situasjonen krever: "... she does it for the right reason – because she understands that this is the right thing to do. And she does this dispositionally – she has a character such that she understands on each occasion what the right thing to do is" (Annas 2003: 8). Har Annas rett i sin kritikk her? Gir Doris et feilaktig bilde av dydene? Både ja og nei. Doris er mye mer nyansert i sin beskrivelse av dydene enn den posisjonen Annas mener han inntar skulle tilsi. Innledningsvis i *Lack of Character* og i sin respons i *Philosophy and Phenomenological Research* anerkjenner han at dyder er karaktertrekk som kjennetegnes av en bestemt type rasjonell overveielse hos aktøren¹⁵. Likevel, det er vanskelig å se at Doris' bemerkninger får noen videre konsekvens for hvordan han betrakter dyd i sammenheng med empiriske studier av karaktertrekk i psykologien. Kan det hende at han er for påvirket av psykologiens forståelse av personlighet og karakter til å gi en redegjørelse av dyd som tar det indre perspektivet alvorlig? For å kunne bruke empirien slik han vil, må dydene presenteres som karaktertrekk. Vi har kanskje her har nådd et punkt hvor dydene skiller seg fra karaktertrekk slik de forstås innenfor psykologien. Dydene er i større grad preget av rasjonalitet og selvbestemmelse enn det rene karaktertrekk trenger å være. Annas mener at dydene må forstås som "... the increasing effectiveness of the agents rationality: the dispositions she develops are the result of her

¹⁴ Annas' poeng kommer kanskje aller best frem der hun skriver at Doris betrakter dydene som noe som er fastlagt og utviklet forutfor og uavhengig av aktivitet (Annas 2005: 637). Hun plukker blant annet ut følgende avslørende sitat fra Doris' bok: "Rather than striving to develop characters that will *determine our behavior in ways substantially independent of circumstance*, we should invest more of our energies in attending to the features of our environment that influence behavioral outcomes" (Doris 2002: 146, min kursivering). Som Annas påpeker er dette en grov misforståelse av dydsetikken, men det synes som om dette er en enkelt glipp fra Doris' side fremfor en generell holdning.

¹⁵ Se for eksempel Doris 2002: 16-18, og Doris 2005: 662-663.

deliberations and decisions, *not* of other factors that can be developed in a mindless way.” (Annas 1993: 51). Doris synes å mene at det primære når man betrakter karaktertrekk, eller dyder, er at de er disposisjoner for en viss type adferd. Annas, på den annen side, ser ut til å mene at det å være dydig innebærer en disposisjon til å imøtekomme ulike situasjoner på en bestemt måte, noe som deretter gir utslag i adferd. Vi kan se dette i et allerede sitert utdrag: ”And she does this *dispositionally* – she has a character such that she understands on each occasion what the right thing to do is” (Annas 2003: 8, min kursivering). Den samme forståelsen av dyd kan vi gjenfinne hos Sreenivasan: ”In general terms, the traits in which virtue theory is interested may be conceived as dispositions to respond in a distinctive way *to a certain sort of situation*” (Sreenivasan 2002: 57). Kanskje er det her snakk om to ulike måter å forstå hvordan dyder disponerer. Annas og Sreenivasan mener at dyd disponerer for en viss type resonering, sensitivitet og beslutningstaking. Doris fokuserer på at dyd disponerer for visse handlinger. Doris’ måte å forstå disposisjon på sammensvarer med hvordan disposisjoner fremstilles i situasjonistenes litteratur, men er denne forståelsen kompatibel med dydsetikernes dydsbegrep?

Vi kan nå vende tilbake til den innvendingen jeg omtalte i introduksjonen til kapitlet: Svekkes Doris argumenter fordi situasjonistenes studier utelater betydningen av praktisk rasjonalitet? Gopal Sreenivasan redegjør for hvordan de to tidligere nevnte forståelsene av disposisjon påvirker hvordan man tolker relevansen av situasjonistenes studier. Han tar utgangspunkt i Doris’ formulering av hva vi forventer av en aktør for å tildele denne et karaktertrekk: ”If a person possesses a trait, that person will exhibit trait-relevant behavior in trait-relevant eliciting conditions” (Doris 2002: 16). For å legitimt kunne påstå at noen har et karaktertrekk kreves det ifølge denne formuleringen først en spesifisering av det aktuelle trekket. En slik spesifisering innebærer å gi en redegjørelse både for hvilken adferd og hvilke situasjoner som regnes som relevante for det aktuelle trekket. Betingelsene for relevans settes av de som gjennomfører studien. Skulle deres krav til adferd og situasjon ikke samsvare med hvordan aktørene i studien selv forstår situasjonen vil dette ikke fanges opp så lenge det kun er adferd som observeres. Annas peker på hvilken betydning dette har: Observatøren og aktørens forestilling om hva som i den aktuelle situasjonen er en dydig handling kan være ulike. Det finnes ingen liste over hvilke situasjoner som korresponderer med hvilke dydige handlinger. I enhver situasjon vil det være flere viktige elementer som må tas opp til vurdering og veies mot hverandre. Vi vil i mange situasjoner konstruere og lese situasjonene vi befinner oss i ulikt. Dermed kan det hende at en observatør mislykkes i å

oppfatte en handling som dydig i øyeblikket, men ved ettertanke likevel innse at det nettopp var dyd hun observerte: "Someone who is reliably kind may do a harsh action which we later understand to be kinder than we thought, given the whole background" (Annas 2003: 14). Konsekvensen er at en observatør kan mislykkes i å registrere konsistens i tilfeller hvor det faktisk er på sin plass å gjøre det. Sreenivasan bruker en kjent studie gjennomført av Hugh Hartshorne og Mark A. May som eksempel. Her undersøkte man konsistens i forhold til ærlighet hos barneskolelever. De undersøkte elevenes evne til å utvise ærlighet i tre ulike situasjoner. Ærlighet skulle påvises ved å se om elevene var villige til å stjele penger gjenglemt på en pult i et tomt klasserom, om de ville lyve for å få en medelev ut av trøbbel, og om de ville jukse på en prøve hvis de ikke risikerte å bli oppdaget. Hensikten var å undersøke i hvilken grad ærlighet opptrer som et konsistent trekk på tvers av ulike (men, relevante) situasjoner. Studien viste at korrelasjonen mellom ærlighet i en situasjon med ærlighet i de andre situasjonene kun var på 23 % (Ross og Nisbett 1991: 98). Doris bruker denne studien som belegg for sin egen posisjon, og konkluderer på grunnlag av denne (og flere liknende studier) at: "It's not that behavior in different situations is completely unrelated ... but that the associations are rather fainter than the marked relationships one would expect if behaviour was ordered by robust traits" (Doris 2002: 63). Denne typen studier viser, ifølge Doris, at ordinær adferd ikke er koordinert av robuste trekk. Sreenivasan derimot, uttrykker bekymring for hvorvidt Hartshorne og May-studien kan brukes for å trekke slutninger av Doris' type¹⁶. Situasjonene som forskerne benytter seg av er for tvetydige, og dermed sårbare for at aktøren og observatøren konstruerer situasjonene ulikt. Kanskje en av elevene ikke tenker på det å plukke opp en slump vekslepenger som sitt eget som å stjele (og dermed ikke som noe uærlig)¹⁷. Dermed er det fra elevens side ikke noen mangel på konsistens mellom denne situasjonen og for eksempel en annen situasjon hvor hun velger å unnlate å jukse på en prøve (Sreenivasan 2002: 58). For Hartshorne og May, derimot, er ærlighet å la pengene ligge, og det er deres tolkninger av situasjonene som ligger

¹⁶ Doris derimot nøyer seg med en bekymring om hvorvidt bruken av barn i studien er til hinder for slutninger om konsistens fordi barn, i mindre grad enn voksne, har "stivnet i formene". Man kan dermed betvile overføringsverdien fra barns til voksnes adferd. Doris legger imidlertid denne bekymringen til side fordi han finner nok grunnlag i studier av voksnes adferd til å legitimere sine konklusjoner (Doris 2002: 63).

¹⁷ Sreenivasan omtaler også hvordan situasjonene synes å være ulike med henhold til relevans for ærlighet. Å jukse på en prøve synes å i mye større grad være et paradigmatisk eksempel på uærlighet enn å beholde gjenglemte vekslepenger. Burde man kanskje benytte seg av situasjoner som har lik grad av relevans? (Sreenivasan 2002: 59).

til grunn for om aktøren regnes som ærlig eller ikke. Gitt disse objektive målestokkene for adferd vil man dermed gå utenom de grunnene som avgjør en aktørs handlinger. For med sikkerhet å kunne påstå noe om adferd og konsistens må man på forhånd forsikre seg om at observatør og aktør deler samme forståelse av situasjonene og trekkene det skal testes for. Dette sitatet oppsummerer innvendingen fint:

... failure to predict a person's behaviour on the basis of 'objective' behavioural measures – that is, low consistency coefficients between such measures – is not always good evidence that the person's behaviour is actually inconsistent across the situations in question. It also depends on the extent to which the subject and the observer agree about the relevance of the trait specifications which the behavioural measures incorporate (Sreenivasan 2002: 58)

Ren observasjon fremstår dermed som en dårlig metode når man skal teste for dyder (og kanskje også karaktertrekk generelt). Et sentralt aspekt, nemlig det indre, fanges ikke opp.

En annen beslektet innvending går ut på at selv om vi har kommet til enighet om en konkret beskrivelse av handling som konstituerer ærlighet, er det ikke gitt at denne beskrivelsen bevarer sin relevans for det aktuelle trekket i alle situasjoner. Ta for eksempel situasjonen hvor en elev lyver for å hjelpe en annen. På forhånd har vi kommet frem til følgende spesifisering av uærlighet: ”Å med vilje fortelle en usannhet i den hensikt å lure en annen”. Gitt en slik definisjon ser Hartshorne og May ut til å lykkes i å plukke ut et paradigmatiske tilfelle av uærlighet. Problemet oppstår her fordi løgnen fortelles for å oppnå et gode, nemlig å få en medelev ut av trøbbel. Vi antar at i denne situasjonen er dette en grunn som veier tyngre enn grunnen for ikke å lyve. Det er altså ikke grunnen som taler for å unnlate å lyve som er den avgjørende grunnen i den aktuelle situasjonen. Det er andre grunner som veier tyngre. Sreenivasan påpeker, hvis dette skal være et paradigmatiske eksempel på uærlighet bør grunnene til ikke å lyve være tilstede i situasjonen som avgjørende for handling (Sreenivasan 2002: 60). Med andre ord, det er ikke åpenbart at den grunnen som bør veie tyngst i denne situasjonen er påbudet om ikke å lyve. Dermed konkluderer Sreenivasan at denne situasjonen ikke egnet til å måle ærlighet, med mindre man mener at for å være ærlig må man fortelle sannheten også i de tilfellene hvor tyngden av grunner taler for å lyve. Vi vil kanskje mene at en som svarer feilaktig for å redde et medmenneske fra en forferdelig

skjebne lyver¹⁸, men jeg vil anta at de færreste vil mene at dette er et velegnet eksempel i en testsituasjon.

Lærdommen man kan trekke av dette er en relativt hverdagslig erkjennelse. Nemlig at vi ikke fullt ut forstår en aktørs handlinger så lenge vi ikke har kjennskap til aktørens grunner (Annas 2003: 14). Det er interessant at i variasjoner over Hartshorne og Mays studie har man funnet at konsistens i adferd i større grad kan påvises hvis testsituasjonene betraktes av deltagerne som beslektede eller å være av samme type. Høyere grad av konsistens finner vi også i de tilfellene hvor trekkene som studeres oppfattes som sentrale i deltagerens selvbylde (Flanagan 1991: 291, Sreenivasan 2002: 65-66). Er ikke dydene nettopp trekk av typen man identifiserer seg sterkt med, som gjør at man er i stand til å plukke ut elementer i ulike situasjoner og forstå at disse ulike elementene påkrever samme dyd? Igjen, økt fokus på indre strukturer gjør at vi kan finne konsistens der man ved hjelp av objektive målestokker mislykkes. Man kan spørre seg hvorvidt det i det hele tatt er mulig å konstruere forsøk som kan fange opp alle sider ved dydene? Det er ikke nødvendigvis tilfellet at situasjonistenes studier er av dårlig kvalitet, de lykkes i å ramme sitt mål, nemlig personlighetspsykologiens forståelse av karakter. Problemet oppstår idet man bruker disse forsøkene for å underminere dydsetikernes særegne forståelse av karakter (Sreenivasan 2002: 64). Må man forstå dydene på Doris' måte for å få de til å passe inn i psykologiens studier? Det er i alle fall ikke åpenbart eller uproblematisk å bruke studier av psykologiens forståelse av personlighet som utgangspunkt for å kritisere en forståelse av karaktertrekk som er basert på filosofiske betraktninger.

På den annen side er det kanskje en svakhet ved dydsetikken at den i liten grad er i stand til å koble sammen konkrete situasjoner med konkrete handlinger, eller å spesifisere hvordan praktisk rasjonalitet munner ut i konkret handling. De vektlegger i stor grad indre strukturer som er vanskelige å påvise. Dette gjør at de kan "sikre" sin posisjon, men prisen de betaler for dette er å svekke teoriens empiriske innhold. Doris mener en slik taktikk rammer teoriens grep om oss som moralske aktører, og dette er jo nettopp dette faste grepet om aktøren som er lukrativt ved dydsetikken. Doris skriver at å gjøre teorien mer empirisk beskjedent "... seems to me to reduce the power of characterological moral psychology as an explanatory

¹⁸ Hartshorne og Mays eksempel er kanskje noe mer trivielt, og spørsmålet om det er rett å lyve i deres eksperiment mer omstridt. Uklarheter til tross, innvendingen holder fordi den baserer seg på å vise at hva som er rett handling i denne situasjonen ikke er ukontroversielt.

and predictive theory, since the theory does not cover the preponderance of subjects” (Doris 2005: 666). Annas og Sreenivasan gjør dyd og dydig handling til noe unødvendig mystisk. Vi er alle kun observatører av andres handlinger, også utenfor forskningssituasjonene mangler vi tilgang til aktørenes grunner¹⁹. Å vurdere om en handling er dydig vil kanskje bli umulig hvis det ikke lar seg bestemme utfra informasjon om observerbar adferd alene. På den annen side, i hverdagen omgås vi i stor grad personer og situasjoner vi er kjent med, og som vi allerede har kunnskap og (velbegrunnede) meninger om. Kanskje kan vi ikke fullt ut anerkjenne betydningen av karaktertrekk så lenge vi befinner oss i en forskningssituasjon. Owen Flanagan minner om at karaktertrekk i stor grad handler om å enklest mulig forholde seg til medmenneskers handlinger: det er enklere å tenke på Per som tålmodig, fremfor å huske hvordan Per handlet i konkrete situasjoner. ”In informationally richer contexts, trait ascriptions are anchored to particular persons and situations and in this way gain rich meaning and circumscription” (Flanagan 1991: 280). I boken *Character* (1991) legger Joel Kupperman frem et forsvar av bruk av litteratur, filosofi og personlig erfaring for å komme frem til moralpsykologiske innsikter. Kupperman mener at ettersom empirisk vitenskap må bedrives i samsvar med strenge metodologiske kjøreregler legger dette visse begrensninger for hvilke typer uttalelser man kan tillate seg å komme med. Kanskje viser dydsbegrepet hen mot grensene for hva som lar seg fange opp vitenskapelig? Er det slik at visse ueliminerbare sider ved moralen prinsipielt sett ikke lar seg empirisk teste?

Det er noe med de redegjørelsene for moralpsykologi som springer ut av filosofi eller litteratur (disse må man anta at er basert på legmenns kjennskap til de psykologiske kategoriene) som gjør at de oppleves med en spesiell klangbunn i oss, og som særskilt viktige: ”... it has to be said that relatively few scientific investigations of moral psychology in recent years have seemed as interesting and important as the results that are found throughout the work of La Rochefoucauld, Hume, and Nietzsche” (Kupperman 1991: 170). Poenget her er at psykologien gjerne konsentrerer seg om å påpeke statistiske sammenhenger og anomalier. I møte med påstander hvor empiri og moral kobles sammen vil den, slik vi har sett, få problemer med å gi en fullverdig redegjørelse den normative komponenten. Psykologer kan forsøke å komme med en deskriptiv redegjørelse for det gode

¹⁹ Med mindre vi ikke eksplisitt spør våre medmennesker om hvorfor de gjorde slik de gjorde, og også i slike tilfeller vil man ikke være sikre på at svarene man får er sannferdige. Vi er alle kjent med fenomener som etterrasjonaliseringer, eller at man vil presentere seg som mer dydig enn man er for egen vinnings skyld.

ved å definere det som hva en gitt gruppe mennesker ville enes om (en metode vi kjenner fra Rawls), men det er ikke gitt at alle moralfilosofiske utsagn lar seg spesifiseres rent deskriptivt (Kupperman 1991: 169). Dydene her kanskje et eksempel på dette. Statistikk er vel og bra, men etikk og moral krever noe mer. Min mistanke er at moralfilosofiske påstander trenger en bestemt form for redegjørelse for å få forfeste hos oss som aktører. Filosofien og litteraturen kan ”pakke inn” sine innsikter på en annen måte enn hva den empiriske vitenskapen kan, en måte som oppfordrer til selvransakelse og refleksjon (Kupperman 1991: 169). Presentert slik kan etikken gjøre det den skal, nemlig å gjøre oss til moralsk sett bedre aktører. Bør moralfilosofien rett og slett være litt spekulativ? Dette bringer oss tilbake til det som ble tatt opp innledningsvis i dette kapittelet, nemlig hvor generell og abstrakt moralfilosofien skal tillate seg å være. Annas’ og Sreenivasans respons på Doris’ utfordring viser at et sentralt aspekt ved dyd vanskelig lar seg teste empirisk, men er dette grunn nok til å forkaste dette aspektet? Nei, jeg vil tro at de indre strukturene som er en del av en dydig aktør er helt sentrale for at vi skal forstå hva dyd er og hvorfor det regnes som et personlig gode å være dydig. For at jeg skal motiveres til å være dydig trengs muligens et slikt ”spekulativt” element. Frem til nå har jeg presisert at moralteorier bør ha best mulig empirisk belegg fordi dette gjør den gjenkjennbar og praktisk relevant. På den annen side: motivasjon sikres ikke utelukkende ved å lage teorier som er mest mulig i samsvar med empiri. Moralfilosofien må tillate seg å heve seg over det empiriske i visse tilfeller. Dette gjør det desto viktigere for filosofer som bedriver moralpsykologi å utvise varsomhet. Kupperman påpeker at filosofisk moralpsykologi ikke har verifikasjonskriterier slik den empiriske vitenskapen har. Det er med andre ord vanskelig å skille de holdbare påstandene fra det som er filosofisk svada. Dette fordrer kritiske lesere: ”This places a burden on the reader, who must perform some of the validation that in scientific work would be part of what is provided; of course, the reader’s validation itself is hardly likely to meet scientific standards” (Kupperman 1991: 170). Kanskje kan man her la filosofi og vitenskap møtes ved å kreve at moralpsykologiske utsagn skal kunne leve opp til visse empiriske standarder. Man må likevel gi rom for at filosofiske utsagn ikke alltid vil la seg testes empirisk. Å bedrive en filosofisk og empirisk sensitiv moralpsykologi vil dermed alltid være en balansegang mellom to ytterpunkter: vitenskapens statistikk som til en viss grad er moralsk irrelevant og filosofiske redegjørelser som kan oppleves som fjerne fra hvordan menneskelivet faktisk fungerer. Man bør forsøke å komme med teorier som er mest mulig åpne for empirisk testing, samtidig som man ikke må forlate noe det som er filosofiens egenart: Å kunne tillate seg å være abstrakt og å generalisere.

2.5 Arpalys respons

Annas og Sreenivasans respons til Doris og situasjonistene bestod i å problematisere hvorvidt dydene kan komme til sin rett i forskningssituasjoner. De fastholder dermed på den tradisjonelle forståelsen av dyder. Nomy Arpaly anerkjenner derimot at situasjonismen problematiserer dydene i så stor grad at vi kanskje må se etter alternativer. Det er her snakk om å forsøke å finne en måte å betrakte dydene som både gjør at dyd fremstår som noe moralsk relevant og empirisk holdbart. Arpalys respons til Doris baserer seg på at karaktertrekk i en hverdagslig kontekst ikke utelukkende dreier seg om å forutse andres handlinger, men om å gi mening til andres og egen adferd. Arpaly ser at dydene, som robuste og globale trekk, er problematiske. Det innebærer dermed ikke, mener Arpaly, at alt snakk om karakter må forlates. Det finnes andre måter å betrakte karakter på som har større holdbarhet. Hun åpner artikkelen med å si seg enig i Doris' avvisning av karakter, og hans alternative teori om karaktertrekk som lokale²⁰: Vi kan kanskje ikke beskrive aktører med de store og generelle valørene slik som ærlighet og tapperhet, eller "personal-ad style" som Arpaly kaller det: "... even if no such thing as the virtue of honesty or the vice of greed exists, it may still be true, for example, that there are such things as John's character or Julia's character, and perhaps such characters cannot be described personal-ad style but could be described in 30 pages by Balzac" (Arpaly 2005: 643). Vi finner ingen enkle regelmessigheter i menneskelig adferd, men man kan kanskje finne regelmessighet hvis man bruker et mer komplekst språk for å beskrive karakter. Arpaly åpner opp for at ønsker og verdier kan betraktes som bestanddeler av en aktørs karakter, og være av betydning for adferd. Vedvarende ønsker og verdier vi identifiserer oss sterkt med har betydning for hvordan vi handler. Vi er ikke tomme skall, og ulik adferd kan ikke kun tilskrives ulike situasjoner.

Arpalys poeng er at det finnes en form for minimal forståelse av karakter som overlever situasjonistenes angrep. Dessverre er ikke dette nok for å kunne gjenopprette dydenes empiriske holdbarhet. At en aktør setter visse verdier høyt kan gi utslag i adferd, men ikke

²⁰ I *Lack of Character* presenterer Doris en måte å (re)konstruere personlighetsstrukturer i lys av situasjonismens funn. Karaktertrekk beskrives her på en måte som ikke innebærer konsistens på tvers av situasjoner, men heller stabilitet gjennom serier med like situasjoner. Vi kan bare med rettmessighet forutse adferd i tilnærmet like situasjoner. Trekk er altså lokale, fremfor globale (Doris 2002: 25).

på en måte som utviser tilstrekkelig forutsigbarhet. Veien fra verdi til adferd lar seg for lett forstyrre av "fremmedelementer", slik som andre mentale tilstander eller aktørens omgivelser. Arpaly bringer likevel inn interesser og verdier fordi dyd gjerne omtales med dette som fokus, fremfor handlingers konsistens eller forutsigbarhet (Arpaly 2005: 645). Det er ikke først og fremst den dydige forutsigbarhet eller stabilitet vi beundrer, men de gode grunnene hun handler på. Forutsigbarhet og konsistens oppstår som en konsekvens av at man har disse grunnene. Den modige "... is more praiseworthy not because he is more reliable, or more predictable, but because his concern for doing the right thing is so deep that the thought of disgrace pains him more than the thought of death" (Arpaly 2005: 646).

Problemet, mener Arpaly, er å anta at det er en enkel korrelasjon mellom grad av verdsetting og observerbar adferd, slik Aristoteles og dydsetikerne gjør. De empiriske studiene Doris bruker viser at en slik antagelse er feilaktig. Verdier og interesser er likevel ikke helt uten forklaringskraft. Selv om vi ikke kan finne noen forutsigbar eller enkelt ordnet korrelering mellom verdier og handling, vil det i visse tilfeller være sant at det å besitte bestemte verdier spiller en kausal rolle og kan forklare handling. I enkeltsituasjoner vil vi sjeldent lykkes i å forutse adferd på grunnlag av slike parametere, men hvis vi betrakter en aktørs adferd over lengre tid, altså summen av en rekke ulike handlinger, vil vi kunne avdekke mønstre. Regelmessighetene vi kan finne vil kunne forklares ut fra verdier og ønsker, og disse vil kunne la oss komme med forutsigelser om en aktørs generelle handlingsmønstre (Railton 2009: 24). Det vi her ender opp med er ikke et argument for dyder, men heller et argument for at vi ikke er fullstendig karakterløse. Dette er Arpaly klar over, og Doris følger henne i dette. Å mangle karakter er ikke det samme som total mangel på innhold. Man er ikke som blanke tavler eller tomme beholdere som kun gis innhold av situasjoner, slik tittelen på Doris' bok kanskje kan gi oss grunn til å tro. Doris peker på fraværet av en bestemt form for karakter, nemlig en som er global (og dermed konsistent) (Doris 2005: 667). For eksempel er det sant at jeg verdsetter miljøvern, og at dette noen ganger gir utslag i handling, men sannsynligvis er jeg ikke konsistent nok i mine handlinger til å kunne opphøye denne verdsettingen til et robust karaktertrekk, selv om det selvsagt sier noe om hvem jeg er: "Deep concerns ... may be profoundly implicated in who we are and how our lives go, so they are intensely important to our intimates, our acquaintances, and ourselves. This remains true when we agree, as Arpaly and I do, that these concerns will not neatly and reliably order behavior" (Doris 2005: 667).

En slik forståelse av karakter er en avstikker fra hvordan man innenfor filosofien tradisjonelt sett har betraktet karakter. Det tillater oss å snakke om karakter, men ikke om dyd. Likevel, i Arpals konklusjon vil dydsetikeren kanskje finne et lyspunkt. Selv om aristotelisk moralpsykologi ikke strekker til hva angår konsistens og forutsigbarhet, kan det kanskje finnes en viktig innsikt hos Aristoteles som handler om å vektlegge betydningen av interesser og ønsker for handlingers moralske verdi (Arpaly 2005: 646). Doris utdyper dette i sitt svar til Arpaly, og her anerkjenner Doris også noe av noe av Annas' og Sreenivasan' kritikk, altså begrensningene ved kun å fokusere på observerbar adferd: "... perhaps my conception of the empirical, with its focus on overt behavior, is unacceptably narrow" (Doris 2005: 669). Å kun anerkjenne observerbar adferd som akseptbar data, slik Doris gjør, kan være en forståelse som er for snever. Av Arpals artikkel fremkommer det at selvforståelse og identitet er aspekter ved moralsk aktørskap som det kan være på sin plass å ta med i betraktningen hvis man ønsker en adekvat moralpsykologi. Å mene at ønskene og verdiene som ligger til grunn for våre handlinger, og hvor dypt disse er rotfestet i bevisstheten, har en betydning for hvordan vi vurderer handlingers moralske status synes plausibelt. Doris mener disse dypt fundamenterte aspektene ved karakteren er sentrale elementer i våre livsfortellinger, og dermed også spiller en viktig rolle for hvordan vi fortolker og former livene våre. *Opplevelsen* av det moralske skjer gjennom vår historieskriving om oss selv. Denne opplevelsen må Doris innrømme at er empirisk, til tross for at den vanskelig lar seg gjøre til objekt for psykologiske studier. En moralpsykologi som er mer realistisk (eller mer livaktig, som Doris skriver) må kanskje innlemme disse aspektene ved moralsk aktørskap.

Ett siste tema i Arpals kommentar er et mer generelt poeng om motivasjon. Det har sitt utspring i studier som viser at tilsynelatende trivielle faktorer i omgivelsene kan ha stor påvirkning for hvordan vi handler. Studiene viser at det å finne en gjenglemte mynt i telefonkiosken eller å ha en kopp med noe varmt i handen, faktorer som vi gjerne mener er bagatellmessige og irrelevante for valgene vi tar, faktisk er svært avgjørende. Vi er mer tilbøyelige til å være hjelpsomme i situasjoner hvor koppen vi holder i er fylt med te fremfor isvann. Jeg kan lære meg at jeg må være mer på vakt ovenfor hvordan jeg handler når jeg har det travelt, fordi jeg vet at å ha dårlig tid kan ha en negativ innvirkning på min tilbøyelighet til å være hjelpsom. Men hvordan skal jeg kunne ta høyde for at det kan finnes andre, ukjente og tilsynelatende betydningsløse gjenstander eller hendelser som kan påvirke hvordan jeg handler? Det finnes sannsynligvis en mengde slike ikke-substansielle elementer vi ikke kjenner til:

The crucial observation is not that mood influences behavior – no surprise there – but just how unobtrusive the stimuli that induce the determinative moods can be. Finding a bit of change is something one would hardly bother to remark on in describing one's day, yet it makes the difference between helping and not (Doris 2002: 30)

En dydsetiker kan innvende at å være dydig innebærer å være sensitiv i forhold til de enkelte situasjonene, og også ta hensyn til og innlemme situasjonens elementer i handlingsoverveieelse. I forhold til å handle under tids-, eller gruppepress synes dette relativt uproblematisk. At dette er faktorer som er viktige for hva vi velger å gjøre er åpenbart. Mer problematisk er det at det ser ut til å finnes elementer som påvirker oss i det skjulte. Vi er ikke i stand til å utvikle en sensibilitet som gjør at vi kan gjenkjenne disse elementene i enhver situasjon. I mangel av en slik evne vil vi kanskje ikke være i stand til å utføre operasjoner som lever opp til dydsetikkens standarder for praktisk overveieelse. Vårt sett med motiver er rett og slett for lite gjennomsluktig. Dette er et sentralt problem for flere versjoner av filosofisk moralpsykologi. Moralsk motivasjon ser i enkelte tilfeller ut til å befinne seg hinsides det området vi har kontroll over. Jeg vil undersøke denne problematikken i større detalj i sammenheng med Robert C. Solomons svar til Doris.

2.6 Solomons respons

Solomons hovedinnvending angår hvorvidt vi kan klare oss uten et begrep om robust karakter. Er vår moralske praksis så tett knyttet opp mot en bestemt forståelse av karakter at å avvise en slik forståelse vil være undergravende for moralen? Fremfor å innta Doris' og situasjonistenes tredjepersonsperspektiv, velger Solomon å betrakte karakterproblematikken i første og andreperson. Dermed bekymrer han seg for hvilke konsekvenser det får for hvordan vi betrakter oss selv når vi beskrives i lys av statistikk og kausale modeller fremfor forståelser som bygger på ansvar, aktørskap og karakter (Solomon 2005: 651). Hva skjer idet vi anerkjenner karakterens kausale impotens og situasjonens kraft? Vil vi ved å anerkjenne at vi er i situasjonenes vold miste all motivasjon for moralsk forbedring og, slik Solomon uttrykker det, gi så godt som hvem som helst en unnskyldning for nesten hva det skulle være? (Solomon 2005: 651) Vil ikke ansvar og aktørskap slik vi kjenner det forvitte hvis vi avstår å betrakte oss selv som (potensielt) utstyrt med om robuste karakterer? Solomon mener at noe viktig går tapt hvis vi ikke lenger gir karakter en sentral plass i

etikken, men heller fokuserer på hvordan ulike former for situasjonelt press fremmer eller er til hinder for gode handlinger. Vi mister tanken om at situasjonenes press er noe vi burde kjempe mot, eller at vi i det minste burde forsøke å forbedre oss slik at vi er bedre rustet til å handle på tross av situasjonene, ikke på grunn av dem: "It is the idea that a person can and should resist those pressures, even at considerable cost to oneself, depending on the severity of the situation and circumstances" (Solomon 2005: 650). Solomon mener altså at karakter er et vern i omgivelser som er moralsk sett risikable og helt sentralt for moralsk forbedring. Å avvise karakter som sentralt etisk begrep vil være undergravende for moralen, og ikke fremme hva man med Doris' ord kan kalle "a healthy first-personal ethical regard" (Doris 2005: 672).

Doris har tre punkter han tar opp i sammenheng med de bekymringene Solomon her lufter. For det første er karakter, forstått som robust og global, et kulturelt sett lokalt fenomen (Doris 2005: 671). Det vil si, det er i den vestlige tradisjonen at karakter fremheves som spesielt sentralt og tillegges stor forklaringskraft innenfor moralsk praksis. Doris plukker ut flere studier som viser at man kan finne kulturer hvor henvisninger til karakter oppleves som mindre relevante. For eksempel ser japanere ut til å i større grad være situasjonister i praksis enn det vestlige er. De er mer tilbøyelige til å tolke adferd i lys av omgivelser og situasjon, fremfor indre disposisjoner (Doris 2002: 105). Det ser derimot ikke ut til at Japan lider av noen form for moralsk forvitring av den grunn. Det er dermed tvilsomt at karakter er uunnværlig for et "sunt moralsk selvbylde"²¹. For det andre og tredje utvikler Doris "karakterløse" måter å forstå henholdsvis ansvar og de reaktive emosjonene som gjerne er knyttet opp til moralsk forbedring. Det kan være verdt å stoppe opp å se litt nærmere på akkurat dette punktet fordi det beveger seg inn mot noe av det mest foruroligende ved situasjonismen. Refleksjonsevne og evne til selvkontroll, evner som vi vil mene at er nødvendige for å tildele ansvar, ser ut til å settes ut av spill i situasjonistenes eksperimenter. Vanligvis vil svikt i disse evnene innebære at aktøren har en unnskyldning for ugjerningen, og dermed ikke stilles som moralsk ansvarlig. Problemet er at situasjonismen lærer oss at disse evnene i radikalt større grad enn vi vanligvis antar er ineffektive. Situasjonismen sår

²¹ Robust karakter er i alle fall ikke knyttet opp mot moralen med nødvendighet. Dog kan det hende at karakterperspektivet er så fundamentert i vår vestlige praksis at det i liten grad lar seg fjerne uten vesentlige endringer i hvordan vi betrakter moral: "It is possible, to be sure, that while East Asians can manage without heavy reference to a Western, globalist, conception of self and other, Westerners cannot; the predicament of culture, after all, is one not easily escaped (Doris 2005: 674) (jfr. Williams i kap. 2.3).

tvil om vår *normative kompetanse*, altså vår evne til å opprette et samsvar mellom verdier og ønsker, evaluative forpliktelser som Doris kaller det, og overveielser som fører til at disse forpliktelsene virkeliggjøres. Arpaly så vi, og Doris fulgte henne i dette, bekymret seg for at elementene som setter denne evnen ut av funksjon både finnes i større antall enn vi vanligvis antar og at de ikke gir seg til kjenne på noen åpenbar måte:

Why was I so amorous? Perhaps I'm in love. Or perhaps it was the smell of garlic or the feel of polyester. If situationism is right, we can play this game with every action. Worse, why should we be confident we've played it well? If we survey 1000 situational factors, how can we be sure that number 1001 wasn't doing the work? Our grip on our motivational universe appears alarmingly frail; we may quite frequently be in the dark or dead wrong about why we do what we do. (Doris 2002: 139)

Denne utstrakte mangelen på tilgang til våre egne motiver gjør at tildeling av ansvar blir problematisk. Siden vi aldri kan være sikre på at det ikke er noen forstyrrende situasjonelle elementer tilstede må vi i utgangspunktet alltid forholde oss agnostiske til hvorvidt en aktør er ansvarlig eller ikke. Doris, derimot, mener at vi kan ta situasjonistenes funn alvorlig uten å måtte radikalt endre hvordan vi forholder oss til ansvar. Doris' redegjørelse går ut på at ansvar kun kan tildeles i de tilfellene hvor vi kan identifisere oss med handlingens bakenforliggende motiv. Umiddelbart synes dette merkelig i lys av situasjonistenes studier: de viser oss jo nettopp at vi i stor grad *ikke* identifiserer oss med de egentlige motivene for handlingene våre. Doris løser dette ved å gi en redegjørelse hvor identifisering med motiver kan gjøres kontrafaktisk i etterkant av handlingen (vi ville identifisert oss med motivet som ligger til grunn for handlingen, men fordi handlingen er en vane reflekterer vi ikke over dette i handlingsøyeblikket), eller i form av å kunne plasseres i en narrativ sekvens vi kan identifisere oss med. Mange av situasjonistenes eksperimenter ser dog ut til å avdekke motiver som ikke lar seg innlemme i noen "selvbiografi", de er mer som psykologiske "tics" (Doris 2002: 143). I disse tilfellene mener Doris at vi må se handlingen i et større perspektiv. Vi identifiserer oss kanskje ikke med den aktuelle handlingen, kanskje kan vi heller ikke holdes ansvarlig for den aktuelle "glippen", men vi kan peke på handlinger som førte frem til uhellet. Her kan vi finne noe å identifisere oss med: "... even where a person fails to identify with the callousness that resulted from haste, he might yet embrace having the sort of packed schedule that induces haste" (Doris 2002: 144). Vi kan ansvarliggjøres for å sette oss selv i "moralske høyrisikosituasjoner". På denne måten mener Doris at tildeling av ansvar kan være på sin plass, selv om denne praksisen ikke er fundamentert i

forestillinger om robuste karaktertrekk. Det er dermed ikke utelukket at situasjonismens funn ikke har noen konsekvenser for hvordan vi betrakter ansvar: kanskje bør vi være mer restriktive med å holde aktører ansvarlig?

Et argument mot Doris redegjørelse er at selv om vi ikke alltid har kjennskap til de motivene som faktisk ligger bakenfor våre handlinger, har vi tilgang til de motivene vi vet vi bør handle på. Gitt at det ikke er fullstendig umulig for oss å utføre disse handlingene er vi ansvarlig for å mislykkes i å gjøre dem²². Vi trenger ikke å plukke ut det som egentlig forårsaket handlingen, og deretter gå en kronglete omvei via identifisering og narrative strukturer. En slik innvending rettes mot Doris av Lawrence Blum (se Blum, 2003). Problemet med en innvending av denne typen er at den kanskje er litt for grovkornet. Selv om vi i teorien er i stand til å gjøre det rette kan det hende at det er for mye å kreve at en aktør skal være herdet mot alle elementer som kan påvirke adferd. Det gir kanskje en uforholdsvismessig streng moral. Doris teori har noe for seg i den grad den er mer psykologisk realistisk. Den er mer human fordi den gir rom for feiltagelser og svakheter som er svært vanlige²³. Den gir oss likevel grunn til å forsøke å bedre vår normative kompetanse ved å lete etter de skjulte elementene som påvirker handling, for deretter å fremme en mer effektiv og virksom resonering (Doris 2002: 146).

Hva med moralsk forbedring? Gjennom Doris redegjørelse for ansvar har vi kunnet se at moralsk forbedring kan foregå gjennom lete oss frem til elementer som virker hemmende på etisk refleksjon og selvkontroll. Forbedring trenger ikke å vokse ut av en robust karakter. Men, det er mer som må gjøres rede for, nemlig de moralske emosjonene som gjerne underbygger streben etter moralsk forbedring (Doris 2005: 671). Spørsmålet er hvorvidt "... a properly developed moral-emotional life requires reference to globalist notions of character" (Doris 2002: 155). Ikke overraskende avviser Doris at dette er tilfellet. Doris er spesielt interessert i skam fordi dette gjerne er den mest fremtredende emosjonen i

²² Jeg er noe usikker på hvor langt situasjonistenes påstander rekker: Det vil si, viser de at situasjonen påvirker oss i den grad at noen handlinger er umulige eller bare svært krevende? Intuitivt vil jeg anta at dette er et gradsspørsmål. Situasjoner hvor vi motiveres "i det skjulte" er tilfeller hvor vår evne til overveielse i stor grad settes ut av spill, og andre handlingsalternativer er nærmest umulige.

²³ Doris poengterer likevel at selv om moralsk klanderverdig adferd er svært utbredt gitt visse faktorer er ikke frekvens egnet for å avgjøre spørsmål om ansvar. Fordi umoralsk adferd er utbredt, også i situasjoner som utøver et relativt mildt press, vil vi derfor få et uønsket overskudd av unnskyldninger. Doris mener vi heller bør gå for et mer diffust kriterium om intensiteten til situasjonelt press for å avgjøre hvilke omstendigheter som er formildende (Doris 2002: 135).

karakteretikk. Skam retter seg ikke mot enkelthandlinger, men har en bredere utstrekning og retter seg mot hvem vi er. Det er selvet som helhet som anklages: "Shame is character's emotional proctor; if our character is unsound, the unpleasantness of shame may spur us to improve it" (Doris 2002: 156). En påstand om at skam på denne måten er nødvendig for å gjøre oss til bedre aktører vil Doris til livs. Doris mener faktisk at skam er mer til hinder for moralsk forbedring, enn et nødvendig element. Doris trekker frem skyld, som er en selvregulerende emosjon på samme måte som skam. Skyld skiller seg fra skam ved kun å knyttes opp mot enkelthandlinger. Å føle skyld innebærer ingen global fordømming av selvet, men kun fordømming av enkelthandlinger. Å tenke på skyld og enkelthandlinger, fremfor skam og global karakter, fremmer en etisk refleksjon som lettere får praktiske konsekvenser. Skyld leder oppmerksomheten vår mot enkelttilfeller hvor vi kan gjøre endringer og forbedre oss. Skam er sjeldent like praktisk rettet, og ender lett i passiv selvforakt. Dermed synes skyld å være mer motiverende for moralsk forbedring, enn den noe mer uklare karakterbyggingen som knyttes opp mot skam: "Rather than requiring global reconstruction, guilt presses us to address particular failings in particular contexts; rather than trying to become a 'nicer person', for example, one might work on treating one's neighbor better when he chatters inanely over the garden fence (Doris 2002: 167).

En annen bekymring fra Solomons side er hva som skjer med mellommenneskelige forhold hvis vi ikke kan anta at våre nærmeste er utstyrt med en karakter som er til å stole på. Solomon påpeker at det å stole på noen er noe utover kun å forutse hvordan de vil handle. Å stole på noen er intimt knyttet til kjennskap om denne personens karakter, mens å forutse handling ikke trenger å være knyttet opp mot karakter i det hele tatt, det kan for eksempel kun være basert på situasjon: " ... prediction of behavior ... involves what I have called *reliance* as opposed to trust" (Solomon 2005: 648). Tiltro er basert på den kjennskapen vi har til våre nærmeste, og fordi vi mener at de har en robust og stabil karakter. Tiltro synes ikke å kunne la seg erstatte av pålitelighet uten at noe går tapt på veien. Må vårt forhold til våre medmennesker være konstruert på en slik måte? Er de så dypt fundamentert på tiltro og karakterkjennskap at overbevisningen om at et slik grunnlag faktisk finnes er en nødvendighet? Kanskje er troen på en robust karakter en nødvendig vrangforestilling? Ofte stoler vi på våre nærmeste i større grad enn vi burde, og ikke sjeldent leder dette til skuffelse. Doris lufter tanken om at vennskap og andre nære forhold ikke alltid er konstruert med et slikt element av "naivitet". Det er grunn til å tro at vennskap ikke utelukkende bygges på anerkjennelse av den andres styrker, men også på den andres svakheter. Vennskap

er et fellesskap hvor man styrker hverandre, og på denne måten er bedre rustet til å møte en verden som moralsk sett er krevende (Doris 2005: 672). I fravær av en robust karakter kan man kanskje ty til robuste vennskap.

Måten Doris takler Solomons bekymringer på gir oss grunn til å være optimistiske på vegne av kompatibilistens prosjekt. Vi var i første kapittel innom hvordan man skal takle at vitenskapen kan avdekke sider av oss som kan virke undergravende for moralen. Solomon bekymrer seg for hvorvidt situasjonistenes funn ville ha en slik virkning for moralsk praksis. Man kunne tenke seg en form for forvitring av moralen i lys av vår lave normative kompetanse. Doris viser at det er mulig å tolke moralen på en slik måte at vi unngår en slik forvitring. Faktisk gir Doris' redegjørelse oss en direkte grunn til å forsøke å forbedre oss, gjennom å være mer observante for omgivelsenes innvirkning. Psykologisk realistiske teorier må unngå å kreve *for lite* av oss som moralske aktører. Det må gis rom for at vi skal forbedre oss (innenfor hva som er psykologisk mulig). Kanskje bør man operere med et lite gap mellom majoritetens faktiske moralske personlighet, og den man konstruerer i en moralteori?

2.7 Noen lærdommer

Hva kan denne debatten fortelle oss om det å bedrive empirisk sensitiv filosofi? Siden vi allerede har nådd en enighet om at empiri har en relevans for filosofiske moralteorier møter vi ikke innvendinger som er beslektet med den naturalistiske feilslutningen. Ei heller blir naturalisme her betraktet som reduksjonisme, derfor møter vi heller ikke Moores "Open-Question Argument". Gitt det rammeverket vi her opererer med synes det mest sentrale å være frykt for at man skal bli *for* naturvitenskapelig orientert. Vi har sett at bekymringer knyttet vitenskapens objektive perspektiv dukket opp i alle tilbakemeldingene. En hovedutfordring for filosofien synes dermed å være å forsyne dette objektive perspektivet med mening.

Annas og Sreenivasan betvilte relevansen av situasjonistenes studier for dyd som filosofisk og moralsk begrep. De anklaget Doris for å grunnlegge sin kritikk på en forståelse av dyd som kun omhandlet de ytre aspektene, altså de som lett lar seg vitenskapelig teste. Dermed ble det ikke tatt hensyn til et sentralt element for å forstå aktørenes handlinger, nemlig

aktørens egen forståelse av situasjonen, og den særegne resoneringen som ligger bak dydig handling. Sreenivasan viste deretter konkret hvordan en kjent situasjonistisk studie på en rekke områder kunne anklages for å trekke feilaktige slutninger fordi de kun loggfører observerbar adferd. Annas og Sreenivasans innvending her er av en art som er velkjent innenfor psykologien: Hvordan håndtere det faktum at en aktørs subjektive tolkning av en situasjon alltid er en kausal faktor i en handlingssituasjon? Aktørers handlinger kan ikke kun forstås som "input" og "output" til et objektivt fiksert stimulus, mellom finnes alltid et tolkende subjekt. Uttrykt annerledes: " ... it is the situation as construed by the subject that is the true stimulus" (Ross og Nisbett 1991: 11). Det er ikke noe særegent ved dyd som moralsk begrep som gjør at det trues mer enn andre psykologiske størrelser av objektiv observasjon. Siden dette er et generelt problem innen psykologien, og ikke et problem som utelukkende oppstår i møtet mellom filosofi og psykologi, gir det oss håp om at det er et problem som kan håndteres. Og, det er god grunn til å tro at det er nettopp det som gjøres. At man ikke bare kan fokusere på objektive målinger, men også må regne med aktørenes subjektive tolkninger er anerkjent innenfor de fleste felter av psykologien. At noe er subjektivt er ikke det samme som at det er tilfeldig. Det er mulig å peke på systematiske faktorer som påvirker og gir opphav til ulike tolkninger (Ross og Nisbett 1991: 77). Så det synes ikke som om dyder prinsipielt ikke lar seg teste empirisk, men det krever kanskje en kjennskap til hva slags psykologiske fenomener dyder er. Doris kommer kanskje på dette området til kort som empirisk filosof, fordi å plukke ut allerede eksisterende forskning ikke resulterer et datamateriale hvor det er gjort forutsetninger som passer dydene. Det er ikke gitt at filosofiske og normative begrep med enkelhet lar seg plasseres i allerede eksisterende studier av psykologiske fenomener. Empiriske filosofer må dermed velge sitt materiale med omhu.

Utfordringer knyttet til vitenskapens objektive perspektiv kan også relateres til Solomons og Arpalys tekster. Solomon bekymret seg over hvorvidt vitenskapens tredjepersonsperspektiv ville virke undergravende fordi den, i mangel av et bedre ord, fører til en "avmytologisering" (disenchantment) av moralen. Vi trenger kanskje å tenke på aktørskap via karaktertrekk for å fungere som moralske aktører. For å si det litt brutalt: Dyder er å betrakte som en nødvendig "livsløgn". En kompatibilist vil her svare at det ikke er i skapelsen av nødvendige illusjoner filosofiens bidrag til moralen ligger. Vi så også hvordan Doris håndterte denne innvendingen ved å vise til at tanken om dyder er et lokalt (vestlig) fenomen, og at en velfungerende moralsk praksis ikke er avhengig av at vi betrakter oss selv

eller andre som utstyrt med robuste trekk. Doris' rekonstruerte forståelse av moralsk ansvar overlater riktignok mye til vitenskapen. Hvor ligger grensene for hvilke motiver vi har tilgang til? Hvilke situasjoner er mest ødeleggende for vår normative kompetanse? Dette er spørsmål som vi kan komme til klarhet i gjennom å se hva vitenskapen har å si om oss, og hvordan vi fungerer. Fremdeles er det sant at vi ikke kan utlede normative *bør* fra vitenskapens *er*. Det er heller ikke hva som forsøkes gjøres her. Når man teoretiserer om moral kan filosofien innta en "ovenfra og ned"-posisjon. Det vil si: Man begynner med overbygningen. Idealene eller moralens standarder fastsettes uavhengig av vitenskapen. Det vitenskapen kan bidra med er å gi oss kunnskap om hva vi kan forvente av verden, eller hvor enkelt våre idealer lar seg virkeliggjøre. Deretter kan filosofien, i lys av vitenskapens funn, forsøke å avgjøre hvor krevende en realistisk moral kan være. Det gjenstår altså mye arbeid for filosofien, både forutfor og i etterkant av at vitenskapen "har gjort sitt". Vitenskapelig påvirkning handler ikke om å finne nye idealer, men om å finne brukbare idealer.

Vitenskapen kan også, gjennom å gi oss kunnskap om hvordan vi fungerer, peke på at moralsk relevans finnes på steder vi vanligvis ikke leter. For eksempel kan situasjonisten peke på en overdreven individualisme i tradisjonell moralteori. Fremfor å være lukkede og isolerte enheter er vi situasjonelt påvirkelige. Vi bør også anerkjenne at moralsk styrke ikke bare hentes i oss selv, men i omgivelsene og hos andre aktører. Vitenskapen oppfordrer derfor til en rikere moralpsykologi. På tross av alle vanskelighetene dydsetikken til nå har havnet i kan dette punktet vendes til noe positivt. Fordi dydsetikken allerede opererer med en mangfoldig moralpsykologi, fremfor å destillere moralen ned til så få elementer som mulig, er kanskje dydsetikken best utrustet til å imøtekomme en slik vitenskapelig motivert pluralisme.

Siden bekymring rundt vitenskapens objektivitet og tredjepersonsperspektiv fremkommer i alle tre tekstene tyder på at det er *noe* her som ikke lett kan avskrives. Arpaly pekte på hvordan et sentralt perspektiv, nemlig den personlige opplevelsen av moralen, ikke gis noen vekt i empiriske studier. På samme måte som en kantiansk moral kan anklages for å være for "kald og fremmed" (fordi den er psykologisk urealistisk) vil en *for* vitenskapelig moral kunne anklages for det samme. Moralens sentrale element i det levde liv er utvilsomt viktig. Filosofi er meningsskapende i større grad enn hva vitenskapen er. Vi husker dette fra Arpalys respons. Opplevelsen av det moralske liv er også et element å ta hensyn til hvis vi vil ha en moralteori som er psykologisk realistisk. Å forene vitenskapelige perspektiver med mer subjektive perspektiver i en enhetlig teori er ingen liten oppgave. Det er kanskje heller

ikke en jobb filosofien skal gjøre alene²⁴, men det viser oss igjen at vitenskapen i seg selv er utilstrekkelig for teoribygging i moralen.

Det er også et ytterligere aspekt ved Annas' og Sreenivasans respons som peker hen mot et dypere, mer filosofisk problem. I møte med Doris' empiriske utfordringer søker Annas og Sreenivasan tilflukt i aktørenes indre strukturer. Det ser nærmest ut som om dyd handler om å tolke situasjoner, fremfor å handle. Det er risikabelt å legge så stor vekt på indre strukturer fordi dydsetikernes påstander om disse strukturene i stor grad bygger på intuisjoner.

Dydsetikken henter mye av sin styrke i å ha en viss intuitiv appell for oss. Dette, har vi sett, bør være et krav til moralteorier generelt, nemlig å nyte en gjenkjennelse fra oss som aktører. Å appellere til intuisjoner er dermed ikke et onde i seg selv, men til en viss grad en nødvendighet: "... intuitions represents the lived phenomenology of ethical life, and whatever status intuitions are ultimately assigned in an ethical perspective, any perspective that ignores them risks distortion and sterility" (Doris 2002: 8). Gjennom denne oppgaven fremheves det at kunnskap om hvem vi er står sentralt når vi teoretiserer om moral.

Intuisjoner er viktige kilder til slik kunnskap, og det bør finnes plass for intuisjoner i teorier om moral. Likevel, etter å ha lest Prinz er vi også klar over at intuisjoner som filosofisk "råmateriale" er kontroversielt. For å kunne argumentere for generelle slutninger på grunnlag av intuisjoner trenger man en forsterkning av disse intuisjonene i form av en ytre bekreftelse på at intuisjonene ikke er illusjoner eller av en svært lokal art. Doris påpeker dette: "It's just that intuition pumps are an obvious instance where the philosophical method is highly speculative. Rather than doing away with reflection on intuitions, I urge augmenting such speculations with less speculative methodologies" (Doris 2002: 9). Det er derfor problematisk at Annas og Sreenivasan motstrider seg en slik ytre bekreftelse, eller i dette tilfellet ugyldiggjøring, av intuisjonene de fundamenterer teorien på. Hvis man vil bruke intuisjoner som elementer i en teori bør man kunne supplere disse med andre metoder som bekrefter at intuisjonene er korrekte. I visse tilfeller vil man da oppdage at en intuisjon er av en tvilsom art, slik som vi har sett med dydsetikken, og man bør dermed ta konsekvensen av dette. Dessuten kan en slik unnvikelse fra empiriske forpliktelser virke mot sin hensikt. Dydsetikken finner sin styrke i å ha en intuitiv appell. Den streber etter å være i overensstemmelse med tanker vi alle gjør oss som moralske aktører. Dette grunnlaget faller

²⁴ Litteratur har ofte blitt nevnt som et viktig supplement til teori i moralske anliggender.

bort når vi oppdager at våre intuisjoner om dyder er av en tvilsom opprinnelse. Annas' og Sreenivasans respons er hjelper dermed lite for å bevare dydsetikken som en akseptierbar teori.

Det er en innvending mot Doris' prosjekt som jeg ikke har tatt med i denne redegjørelsen. En dydsetiker vil kunne hevde at de psykologiske studiene er i samsvar med hva dydsetikken påstår, nemlig at det finnes svært få som er fullt ut dydige. At de færreste av oss lever opp til dydsetikkens absolutte ideal er å forvente. Doris anerkjenner en slik innvending. Han kan ikke avvise at det er en mulighet for eksepsjonelle aktører som er i besittelse av robuste karaktertrekk. Problemet for dydsetikken er at disse sjeldne individene utgjør et avvik fra det normale. Vi har god grunn til å tro at å kultivere dyder ikke er en naturlig væremåte for oss. Det er nettopp i en slik henvisning til naturen at dydsetikken henter et sentralt argument for sin teori: nemlig, at de har sitt utgangspunkt i betraktninger rundt hvordan vi er fra naturens side. I neste kapittel skal vi se hvordan betraktninger om menneskelig natur medfører å anerkjenne naturalismens grenser. Menneskene er føyelige og foranderlige. Derfor kan vår naturlige væremåte aldri fikses slik at man kan grunnlegge teorier på hvordan vi er fra naturens side. I normative spørsmål kan dermed naturalismen bare følge oss et stykke på veien.

3. Gjør dydene oss lykkelige? Om naturalismens grenser

Hvor henter dydsetikeren kraften i påstanden om at dydene er psykologiske strukturer det er verdt å utvikle? Eller, hvordan redegjør dydsetikerene for at vi har grunner til å handle i samsvar med dydene? Først og fremst appellerer de til at å være dydig er gunstig for oss. Å være dydig er konstituerende for en bestemt form for objektiv velvære, *eudaimonia*. Eudaimonia, heretter lykke, er ikke ment å være en hvilken som helst form for velvære. Det er en lykke som ikke utelukkende kan knyttes opp mot den enkeltes nytelse, men oppnås gjennom å leve et liv som er i tråd med hva naturen har nedlagt for oss. Det er kun som dydig at man er lykkelig, fordi menneskenaturen er av en slik art at det er kun gjennom dydig handling at vi ”belønnes” med lykke (Annas 2006: 526). Dydenes normative grunnlag hentes fra naturen, i fakta om oss som naturvesener. Dydsetikerne er altså naturalister: ”Neo-Aristotelian kinds of virtue theory claim not only that it benefits me as an individual to be virtuous, but also that it benefits humans to have the virtues because of the kind of animals that we are” (Annas 2006: 527). Gjennom denne oppgaven har det blitt argumentert for at moralsk naturalisme er en vei det verdt å begi seg ut på. Men, i det naturalistiske landskapet finnes også blindveier, og vi har her kommet frem til én. Det er nemlig grunn til å tro at hvis en teori skal være i tråd med hva vitenskapen kan fortelle om hva som er konstituerende for lykke eller vellykkede liv må denne teorien være pluralistisk. Det vil si, den bør gjenspeile alle de forskjellige måtene (vellykkede) menneskeliv lar seg realiseres på. Man bør dermed være varsom med å knytte lykke opp mot en bestemt livsform. Det finnes rett og slett hverken én universell menneskenatur eller én måte å leve på som er naturlig for mennesker. Det hviler en form for vilkårlighet over dydene. Dette truer dydenes normativitet fordi insentivene vi har for å utvikle ett spesifikt sett med dyder er at nettopp denne psykologiske oppbygningen gjør oss i stand til å leve best mulig. Både Flanagan og Prinz tar tak i en slik innvending. Jeg tror denne kritikken er delvis berettiget. Det er mulig å forsvare noen av verdiene som presenteres gjennom dydene, slik som rettferdighet, måtehold, mot og nestekjærlighet. Gjennom å implementere disse verdiene i samfunnet kan vi oppnå menneskelig velferd, både som enkeltindivider og som deltagere i en større sosial enhet. Jeg vil ikke presentere et slikt forsvar her, bare notere meg at det i alle fall ikke synes åpenbart feilaktig. Derimot vil jeg heller konsentrere meg om noe mer problematisk ved dydsetikkens påstander. Problemet er at dydsetikerne mener at disse verdiene må virkeliggjøres gjennom

kultivering av dyder og dydig aktivitet. Vi skal se at dette ikke fremstår som plausibelt hvis vi betrakter samfunn med annen kultur og verdensanskuelse enn den vi har i vesten.

Både Flanagan og Prinz påpeker at vår naturlige utrustning normativt sett er ”a mixed bag” (Flanagan 1991: 51; Prinz 2009: 133). Det er mange egenskaper vi kan sies ha fra naturens side, men dette gjør de ikke gode eller ”lykkefremmende”. Prinz mener vi har vitenskapelig dekning for å hevde at aggressivitet knyttet til beskyttelse av egen eiendom er naturlig. Vi mener likevel ikke at slik aggressivitet er moralsk sett godt. Å fremme en bestemt væremåte fordi den er naturlig forutsetter at mindre tiltalende sider av mennesket også må redegjøres for. Dydsetikeren må skille mellom egenskaper som er gode og de som ikke er det. Dette, mener Prinz, vil de mislykkes i. Det kan ikke gjøres uten å ende i en sirkelargumentasjon hvor man hevder at visse naturlige egenskaper er gode fordi de stemmer overens med dydene, samtidig som man henter dydenes normative kraft fra deres opprinnelse i menneskenaturen (Prinz 2009: 134).

Både Prinz og Flanagan har rett i å påpeke at ”naturlig” ikke er det samme som ”god”, men dydsetikerne trenger ikke å gå i denne fellen. Det er et naturlig faktum om oss (som mennesker) at vi kan stille oss kritisk til egenskaper vi har og vurdere ulike livsformer. At vi kan innta denne kritiske distansen er ikke noe mer enn enda et naturlig faktum om oss, nemlig om vår rasjonalitet. Fordi vi er rasjonelle kan vi evaluere den moralske verdien til egenskaper vi har. Muligheten til å være kritisk til vår natur, er gitt oss av den samme naturen: ”The other animals live 'the way' they do because it is in their nature to do so; we do not. They cannot contemplate alternatives and decide to change things, or choose to try a new way as we can; they are biologically determined, we are not” (Hursthouse 1999: 220). Dermed kan dydsetikerne åpne opp for å regne visse naturlige egenskaper som gode, uten å måtte anerkjenne alle. Dydsetikeren trenger derimot en redegjørelse for lykke eller vellykkethet som kan fungere som kriterium når de vurderer våre egenskaper. Noen naturlige egenskaper er dyder fordi de leder til oss til lykkelige liv. Her møter vi igjen innvendinger fra Prinz og Flanagan: Hva som er lykke eller gode liv for oss mennesker er ikke noe som kan leses av fra naturen. Menneskelig lykke lar seg ikke fastsette, og dermed kan vi heller ikke nedlegge ett sett med dyder som er gode eller fordelaktige for oss: ”One problem is that the idea of the most natural environment conducive to well-being cannot be fixed for persons the same way it can for acorns or orchids. Our natures are too plastic and our potentialities too vast for that” (Flanagan 1991: 51).

Hursthouse er enig med Flanagan og Prinz. Fordi vi, i motsetning til alt annet levende, er rasjonelle, lar menneskelig lykke eller en livsform som er optimal for oss, seg ikke diktere av naturen: "... the idea that nature could be normative with respect to us, that it could determine how we should be, is one we will no longer accept" (Hursthouse 1999: 220). Hursthouse mener at selv om vi ikke determineres av naturen, finnes det en likevel en mulighet for å hevde at noen menneskelige livsformer er bedre enn andre. Det er én egenskap som er karakteristisk menneskelig, og denne egenskapen begrenser hvilke levemåter som kan sies best å ivareta vår velferd. Slik vi så, vår rasjonalitet er karakteristisk for oss som art, og gjør at vi kan ha en kritisk distanse til oss selv som ikke er mulig for andre vesener. Men denne evnen isolerer oss ikke fra resten av naturen. Vi kan studere planter, og avgjøre hva som må være tilstede for at en blomst skal utvikle seg og fungere godt som blomst. Noe av det samme kan vi gjøre for mennesker, selv om vår rasjonalitet ikke tillater samme grad av spesifisitet og detaljrikdom som andre arter. En karakteristisk menneskelig måte å leve på er en rasjonell måte, og dette innebærer: "A 'rational way' is any way that we can rightly see as good, as something we have reason to do. Correspondingly, our characteristic enjoyments are any enjoyments we can rightly see as good, as something we in fact enjoy *and* that reason can rightly endorse" (Hursthouse 1999: 222). En karakteristisk menneskelig aktivitet er altså normativ aktivitet. Vi kan her mistenke at dydsetikerne legger så stor vekt på vår rasjonalitet at naturalismen som var utgangspunktet for å fremme dydene blir et perifert poeng. Dette er ikke tilfellet, mener Hursthouse. Naturalismen står fremdeles sentralt fordi den setter rammer for vår normative virksomhet. Mennesker lever på en måte som fører til lykke kun hvis de tar del i aktiviteter og samfunn som ivaretar individuell overlevelse, artens overlevelse, artskarakteristisk velvære og fravær av smerte, og tilslutt, den sosiale gruppens velfungerenhet (på en artsspesifikk måte) (Hursthouse 1999: 202). Alle disse er kriterier for hvordan vi som mennesker skal fungere godt som biologisk vesen. Altså gjenstår det et betydelig bidrag fra naturen når vi skal bestemme og vurdere ulike livsformer. Å leve på en måte som tilfredsstiller disse kriteriene innebærer å kultivere og handle etter ett sett med dyder. Et slikt levesett er den beste måten å være menneske på, og den måten som gjør oss lykkelige. Hursthouse mener altså at kultivering av dydene og dydig aktivitet er det som fremmer menneskelig velferd. Videre mener hun at vi kan bestemme ett sett med dyder som sikrer denne velferden.

Jeg vil først ta for meg den delen av dydsetikken jeg mener er ”defekt”, nemlig at dydene er den naturlige måten for oss mennesker å leve på, og den måten som gjør oss lykkelige. Hvorfor mene at de fire kriteriene (individuell overlevelse, artens overlevelse, artskaraktistisk velvære og fravær av smerte, og den sosiale gruppens velfungerenhet) som ligger til grunn for at vi skal fungere godt kun kan tilfredsstilles gjennom dydene? Eller uttrykt annerledes: Hvordan kan dydsetikerne hevde at det er en bestemt psykologisk struktur som er naturlig for menneskene, og det kun er gjennom å kultivere denne strukturen at vi ”fungerer godt”? Hursthouse mener at dydene er skreddersydd for å fylle en slik funksjon fordi de innebærer en kultivering av visse aspekter ved oss. Sentralt her er at mennesket betraktes som helhet, og som del av naturen på lik linje med andre levende vesener. På samme måte som det finnes visse aspekter ved en plante som avgjør om den er et godt tilfelle av sin art, finnes det aspekter ved oss mennesker som avgjør om vi er velfungerende. Noen av disse aspektene deler vi med andre arter, men grunnet vår rasjonalitet gis de en litt annen tolkning når de brukes for å vurdere mennesker. Det er i alt fire slike aspekter. De omfatter prosesser i oss som er noe utover det rent fysiske, adferd, emosjoner og ønsker, og rasjonalitet (Hursthouse 1999: 207). Når vi lever gode liv er disse aspektene ”godt kalibrert”. Og veltilpassethet i disse dimensjonene gjør at vi oppfyller kriteriene fra forrige avsnitt, og lever etisk gode liv. Moralsk vurdering vil altså også omfatte evaluering av hvordan den enkeltes tilstand forholder seg i disse aspektene. Vi så i kapittel 2.2 at dydene ikke kun er rasjonelle evner, men også innebærer visse emosjonelle og sansemessige disposisjoner. Å være dydig handler om å fungere godt i flere dimensjoner, også de dimensjonene som menneskene ikke er alene om å befinne seg i. Å være dydig handler også om å kunne sanse noe som etisk viktig²⁵, å ha visse emosjonelle disposisjoner og selvsagt også at det er en overensstemmelse mellom rasjonalitet og handling. Derfor mener Hursthouse at å leve i tråd med dydene er det som fremkommer som den ideelle måten å leve på for oss mennesker. Dydene involverer vurdering og forbedring i alle de fire aspektene: ”Hence the concept of a virtue emerges as apparently tailor-made to encapsulate a favourable evaluation of just those aspects which, according to the naturalism here

²⁵ Derfor henvisningen til prosesser som er noe mer enn kun fysiske. For andre vesener vil dette aspektet tilsvare det å ha god helse, altså utelukkende det fysiske. For mennesker derimot, regner vi det vanligvis ikke som moralsk relevant hvorvidt vi er friske eller ikke. Det er absurd å mene at å bruke briller er en moralsk skavank, eller at etter hvert som vi blir eldre og helsen svikter, blir vi også mer og mer umoralske. Det finnes likevel noen prosesser som i stor grad er fysiske, men hvor noe kognitivt blander seg inn, slik som når den dydige nærmest intuitivt sanser hva som er etisk relevant i omgivelsene. På grunn av dette finner Hursthouse likevel plass til et slikt aspekt hos mennesker (Hursthouse 1999: 208).

outlined, are the ethically relevant ones” (Hursthouse 1999: 208). La oss for øyeblikket se bort fra enkeltverdiene som virkeliggjøres gjennom dydene, og heller betrakte dydene utelukkende strukturelt. Å være dydig innebærer å ha visse reaksjonsmønstre, og å utvise konsistens og stabilitet. Er det kun gjennom å strebe etter en slik oppbygning at verdiene dydene representerer kan virkeliggjøres? Og videre, er det kun gjennom dydene at vi lever moralsk godt, og at vi opplever lykke og velferd på en karakteristisk menneskelig måte? Det er her dydsetikerne trår feil. Det kan godt være at dydsetikerne har rett i å hevde at det finnes et sett naturlige kriterier som ligger til grunn for gode liv: individuell overlevelse, artens overlevelse, artskarakteristisk velvære og fravær av smerte, og tilslutt, den sosiale gruppens velfungerenhet. Dette er naturlige fakta om oss som synes plausible å innlemme i en naturalistisk redegjørelse for moralen. De er dessuten så generelle at de tillater at et vidt spekter av ulike moralsystemer kan tilfredsstille disse kriteriene. Railton reduserer moralen til naturlige egenskaper som synes å være i overensstemmelse med dydsetikernes kriterier. Han skriver om sitt prosjekt: ”... there are perfectly general criteria of moral assessment, nonetheless, by the nature of these criteria no one kind of life is likely to be appropriate for all individuals and no one set of norms appropriate for all societies and all times” (Railton 1986: 165). Dydsetikerens kriterier utgjør kanskje ikke en uttømmende liste. Eller uttrykt på en annen måte, kriteriene er kanskje hverken nødvendige eller tilstrekkelige til å ligge til grunn for gode moralske liv, men de er i alle fall ikke åpenbart feilaktige eller har en iboende defekt.

Påstanden om at dydene er veien å gå for og nå et slikt mål er derimot ikke holdbar. Retter vi blikket mot andre kulturer fremstår dydene som en spesifikt vestlig konstruksjon. Prinz påpeker dette: I lys av andre kulturer har vi grunn til å betvile holdbarheten til påstanden om at lykke eller velferd må være et resultat av kultivering av karaktertrekk (Prinz 2009: 136). Hvis vi, slik vi har sett andre filosofer gjøre gjennom denne oppgaven, henter inn data fra psykologi, sosiologi eller antropologi kan vi betrakte dydsetikken fra et annet perspektiv. Psykologen Richard Nisbett fremhever kontrastene som finnes i verdensbildene til asiater og vestlige:

... modern Asians, like the ancient Chinese, view the world in holistic terms: They see a great deal of the field, especially background events; they are skilled in observing relationships between events; they regard the world as complex and highly changeable and its components as interrelated; they see events as moving in cycles between extremes; and they feel that control over events requires coordination with others.

Modern westerners, like the ancient Greeks, see the world in analytic, atomistic terms; they see objects as discrete and separate from their environments; they see events as moving in linear fashion when they move at all; and they feel themselves to be personally in control of events even when they are not. (Nisbett 2003: 109)

Disse ulike verdensbildene involverer både at man bokstavelig talt sanser verden annerledes, men også at man forholder seg ulikt til omgivelsene på et konseptuelt nivå. Nisbett lister opp noen vesentlige forskjeller: Vestlig tankesett domineres av en ”enten-eller”-måte å tenke på i forhold til logikk og kausalitet. Man forestiller seg gjerne at en hendelse har én årsak fremfor å tenke at årsakssammenhenger kan være mer komplekse, slik asiater har for vane å tenke. Vestlige ser ut til å ha en dyptliggende hang til å oppfatte og plukke ut enkeltfaktorer, fremfor å være mer holistiske i sin verdensanskuelse²⁶. Asiater betrakter verden som i forandring i større grad enn vestlige. Og, speilbildet av dette er at vestlige ser etter stabilitet og konsistens (Nisbett 2003: 203-205). Med en vestlig bakgrunn vil det dermed kanskje være mer nærliggende å verdsette stabile enheter, slik som karaktertrekk og dyder? ”The idea that virtues are necessary for (or even conducive to) flourishing may derive from the Western tendency to value consistent identity over time” (Prinz 2009: 135). Gitt vår vestlige historie, kultur, og samfunnsmessige oppbygning er kanskje dydene viktige for velferd og lykke²⁷, men det ser ikke ut til å være slik alle andre steder. Dyder er ikke noen allmenmenneskelig og naturlig psykologi, og heller ikke noen universell forutsetning for ”det gode liv”. Et eksempel kan hentes fra buddhismen hvor nirvana gjerne regnes som den ultimate lykke (hvis nirvana i det hele tatt lar seg beskrives i slike termer). Å nå nirvana innebærer en fullstendig oppløsning av selvet. Dette står i sterk kontrast til en aristotelisk og vestlig tradisjon hvor lykke oppnås gjennom å være dydig, noe som innebærer en ”herding” av selvet (Prinz 2009: 135).

²⁶ Det er interessant at denne forskjellen ser ut til å kunne forklares av utelukkende kulturelle og sosiale faktorer: Ulike måter å forklare hendelsesforløp på oppstår ikke før tenårene. Slik Nisbett skriver: Det tar tid å lære seg hvordan man skal forklare adferd i tråd med de kulturelle normene (Nisbett 2003: 115). Både vestlige og asiatiske forklaringsmodeller ser ut til å ha en kulturell opprinnelse, fremfor å være naturgitt.

²⁷ Nå har vi riktig nok gode grunner, jfr. Doris’ innvendinger, til å betvile hvorvidt det å være dydig er noen god oppskrift for å sikre velferd også i vesten. Nisbett mener faktisk at dette er et tilfelle hvor vi ikke bare skal akseptere at det er kulturforskjeller, men heller legge oss til en praksis som er nærmere den vi finner i Asia. I Asia anses adferd gjerne som mer kompleks og som resultat av samspill mellom person og situasjon. Dette er et mer korrekt bilde av hvordan adferd genereres (Nisbett 2003: 207).

Vi kan dermed fastslå at dydsetikerne ikke kan hente dydenes normative grunnlag fra naturen. Vi kan ikke hevde at å utvikle dyder er den naturlige veien å gå for å realisere en spesifikk form for menneskelig lykke. Den lykken som dydene er ment å medbringe er ikke mer naturlig enn lykke oppnådd på andre måter. Andre former for velferd eller lykke er ikke av noen dårligere kvalitet, selv om den ikke er et resultat av ervervelse av dyder. Prinz uttrykker det beintfrem: "Faith in a universal form of well-being teeters between cultural chauvinism and a form of naturalism rivaling [Herbert] Spencer's in its naiveté and vulgarity" (Prinz 2009: 136). Uten noe grunnlag for dydene i naturen mister dydene sin normative appell. Dydsetikerne kan kanskje komme med en påstand om at dydene er verdt å kultivere i vestlige samfunn fordi vestlige utsettes for sosialisering i en kultur som fremmer og verdsetter kvaliteter som besittelse av dyder medbringer. Når vi betrakter Doris' kritikk ser kultivering av dyder fremdeles ikke så tiltalende ut. Og når vi i tillegg oppdager at andre kulturer blomstrer uten noe begrep om dyder blir det enda mer problematisk.

Vi har sett gjennom flere kapitler at naturen kan fortelle oss mye nyttig om hvem vi er, og at denne informasjonen kan være med på å forme moralteorier. I dette kapittelet ser vi derimot hva naturalismen *ikke* kan fortelle oss, nemlig en spesifisering av hva som er menneskelig lykke. Realisering av velferd må kunne anses å være moralens endelige mål. Empiriske vitenskaper kan gi oss kunnskap om hvilke forutsetninger vi ikke kan gjøre (fordi forutsetningene er urealiserbare). Kanskje kan de til og med forsyne oss med en form for minimal basis for menneskelig lykke, noen biologiske behov må tilfredsstilles for at vi skal kunne realisere noen form for velferd. Likevel, en god del gjenstår for å se hvordan vi best mulig kan realisere en delt velferd. Dette rommet åpner for kritisk refleksjon om hvem vi er, og hvem vi bør være. Vi kan se til andre kulturer eller andre epoker for å avgjøre hvilke muligheter som er åpne for oss, og i hvilken grad vi kan endre oss. Dette er et fellesmenneskelig prosjekt, men fordrer aktører med ferdigheter som blant annet kultiveres gjennom filosofisk trening. Martha Nussbaum fremhever hvordan opplæring i kunst og humanvitenskapene er nødvendig for å skape deltagere som kan fungere godt som medlemmer av et slikt fellesskap. Det krever nemlig "... an active, critical, reflective, and empathetic member of a community of equals, capable of exchanging ideas on a basis of respect and understanding with people from many different backgrounds" (Nussbaum, 2010). Filosofi er dermed uunnværlig for å bringe moralen videre, der naturalismen ikke kan fungere som fundament for moralen.

Konklusjon

Gjennom denne oppgaven har vi utforsket forholdet mellom empirisk vitenskap og moral med trioen naturalisme, kompatibilisme og psykologisk realisme som rettesnor. Jeg tror at oppgavens redegjørelse for en slik holdning til moralfilosofien best oppsummeres ved å betrakte følgene dette trekløveret har for moralfilosofi i to dimensjoner. Det kan pekes på konsekvenser både for hvordan man driver moralfilosofi, altså metodologisk, og for hvordan moralteorier bør se ut, altså hvilke elementer man opererer med. Et viktig stikkord i begge dimensjonene er *mangfold*, eller *pluralisme*. Til slutt vil jeg gjøre meg noen tanker om hvorfor jeg tror at dette er en lovende utvikling innen moralfilosofien, altså hvorfor jeg tror naturalisme, slik den har blitt presentert gjennom denne oppgaven er ”en filosofi for fremtiden”.

Naturalisme som kompatibilisme innebærer å betrakte kunnskap om hvordan vi er som relevant for konstruksjonen av moralteorier. Slik kunnskap kan ha sitt opphav i intuisjoner eller være fakta hentet fra empiriske vitenskaper. Vi har sett at intuisjoner i filosofi er materiale man må behandle med varsomhet. Intuisjoner er ikke en type kunnskap som kan betraktes isolert fra empirisk vitenskap. Vi kan sette det hele på spissen ved å betrakte tradisjonell moralfilosofi i et litt ufordelaktig lys: En type filosofi som kun baserer seg på intuisjoner produsert fra lenestolen, er ikke holdbar. Det er ikke tilfellet at intuisjoner i seg selv ikke er brukbare som støtte for teori, men henvisninger til intuisjoner kan kun gjøres med rettmessighet hvis man også kan fremlegge argumenter for at intuisjonen er gyldig. Dette forutsetter at man kan si noe om de ulike psykologiske prosessene som produserer intuisjonene. En naturalistisk kompatibilist vil i tillegg ikke nøye seg med å komme med empirisk belegg for intuisjoner. I sin filosofi vil han anse empirisk data som relevant i seg selv, altså som førstehånds belegg for en teori. Han vil konstruere teorier som på best mulig måte er i samsvar med hva kognitive vitenskaper kan fortelle oss om hva slags vesener vi er. En kompatibilist vil altså i stor grad la sin forskning lede av hva andre vitenskaper har å si om emnet som studeres. For kompatibilisten vil det dermed ikke finnes én metode for å generere påstander innen moralfilosofi. Hva vi har er et mangfold av metoder, og kompatibilistens holdning til sitt forskningsfelt kjennetegnes ved en metodisk pluralisme. Filosofien opphører dermed ikke som selvstendig disiplin. Filosofien vi kjennetegnes av en teoretisk holdning, av å være meningsskapende, og gjennom hvilke emner som studeres.

En slik metodisk holdning gir seg til kjenne i hva slags teorier kompatibilisten fremmer. Innlemmelsen av empirisk forskning i moralfilosofi vil ha konsekvenser for teoribygging på to måter: Først kan vi bemerke oss at empiri virker *begrensende*. Med dette mener jeg at empirisk forskning forteller oss noe om hvem vi er, og denne informasjonen setter grenser for hvordan kompatibilistens moralpsykologi kan se ut. Det innebærer at noen velkjente byggesteiner i moralteori, slik som dydene, må gis en annen betydning, eller kanskje til og med forlates. Det kan også føre til at man må anerkjenne den menneskelige bevisstheten som mer mangfoldig enn hva tradisjonelle moralteorier har antatt. Kanskje er ikke moralsk motivasjon av én type, og moralsk overveielse involverer kanskje en rekke komplekse prosesser. Enkelhet er kun et gode hvis empirien tillater det. For det andre: Å være naturvitenskapelig orientert medfører å anerkjenne menneskenes mangfoldighet. Vitenskapen kan fortelle oss noe om hvordan vi ikke er, men åpner også opp et stort rom for hvordan vi *kan* være. Vår natur vil dermed avsløre at det finnes en rekke måter menneskelivet kan realiseres på, og ”gode” liv lar seg ikke utledes av vår natur. Kompatibilisten anerkjenner at det ikke finnes noen universell moral, men en pluralisme av ulike moraler. Dette er et emne som ikke er blitt tematisert i denne oppgaven, men som peker mot veier videre. Nemlig om vi skal erkjenne dette som relativisme, eller hvorvidt det er mulig å unngå en slik posisjon?

Hvorfor mener jeg at man bør være moralsk kompatibilist? Rent metodisk er det ingen vei utenom å anerkjenne at henvise til intuisjoner uten videre belegg for disse er risikabelt. Å åpne opp for andre vitenskaper på dette området synes dermed nødvendig. I hvilken grad man bør innlemme empiri i filosofens arbeid utover dette er kanskje vanskelig å avgjøre med eksakthet. Sannsynligvis vil det være betinget av emnet man studerer, men jeg har gitt gode grunner for å anse empiri som relevant for etikken. Moralfilosofien vil kanskje i stor grad dreie seg om å opprette refleksiv likevekt mellom moralske postulater og empiriske data. Kanskje vil det føre med seg at man returnerer til en moral som er mindre opptatt av det tekniske, begreper og analyse, og mer opptatt av mennesker og mening. Filosofien bringer mening inn i et vitenskapelig verdensbilde, og gjennom å være empirisk sensitiv kan moralteorier oppleves som mer livaktige og relevante for oss som aktører. Dessuten vil det å introdusere en slik dimensjon, altså en naturvitenskapelig, være utfordrende og spennende. Utfordringen og gleden ligger i å skape teorier som er holdbare også i denne dimensjonen, fremfor å skape teorier hvor, slik Richard Nisbett uttrykker det, ”The joy lies in modeling for its own sake, not in making sense of behavior” (Nisbett 2003: 205).

Kildeliste

- Annas, J. (1993). *The Morality of Happiness*. Oxford: Oxford University Press.
- Annas, J. (2001). "Moral Knowledge as Practical Knowledge". *Moral Knowledge*. E. E. Paul, F. D. Miller og J. Paul (red.). Cambridge: Cambridge University Press. 236-256.
- Annas, J. (2003). "Virtue Ethics and Social Psychology". Notat til forelesning holdt ved Mershon Center for International Security Studies. Tilgjengelig online på <https://kb.osu.edu/dspace/bitstream/handle/1811/32006/Virtue%20Ethics%20and%20Social%20Psychology.pdf?sequence=2> (oppsøkt 28.02.11).
- Annas, J. (2005). "Comments on John Doris' *Lack of Character*". *Philosophy and Phenomenological Research*. 71: 636-642.
- Annas, J. (2006). "Virtue Ethics". *The Oxford Handbook of Ethical Theory*. D. Copp (red.). Oxford: Oxford University Press. 515-536.
- Anscombe, G. E. M. (1958). "Modern Moral Philosophy". *Philosophy*. 33: 1-19.
- Arpaly, N. (2005). "Comments on *Lack of Character* by John Doris". *Philosophy and Phenomenological Research*. 71: 643-647.
- Blum, L. (2003). "Review of Doris, *Lack of Character*". *Notre Dame Philosophical Reviews* 02.08.03. Tilgjengelig online på <http://ndpr.nd.edu/review.cfm?id=1344> (oppsøkt 03.04.11).
- Cavell, S. (2010). *Little Did I Know: Excerpts from Memory*. Stanford: Stanford University Press.
- Doris, J. (2002). *Lack of Character: Personality and Moral Behavior*. Cambridge, New York: Cambridge University Press.
- Doris, J. (2005). "Précis of *Lack of Character*". *Philosophy and Phenomenological Research*. 71: 632-635.

-
- Doris, J. (2005). "Replies: Evidence and Sensibility". *Philosophy and Phenomenological Research*. 71: 656-677.
- Feldman, R. (2001). "Naturalized Epistemology". *Stanford Encyclopedia of Philosophy*. Tilgjengelig på <http://plato.stanford.edu/entries/epistemology-naturalized/> (opp søkt 15.05.11)
- Flanagan, O. (1991). *Varieties of Moral Personality: Ethics and Psychological Realism*. Cambridge, Mass: Harvard University Press.
- Flanagan, O., H. Sarkissian & D. Wong (2008). "Naturalizing Ethics". *Moral Psychology: The Evolution of Morality: Adaptations and Innateness*. Bd 1. W. Sinnott-Armstrong (red.). Cambridge, Mass.: MIT Press. 1-25
- Hursthouse, R. (1999). *On Virtue Ethics*. Oxford: Oxford University Press.
- Johnson, M. L. (1996). "How Moral Psychology Changes Moral Theory". *Mind and Morals: Essays on Cognitive Science and Ethics*. L. May, M. Friedman og A. Clark (red.). Cambridge, Mass: MIT Press. 45-68.
- Knobe, J. og S. Nichols (2008). "An Experimental Philosophy Manifesto". *Experimental Philosophy*. J. Knobe og S. Nichols (red.). Oxford: Oxford University Press. 3-14.
- Kupperman, J. (1991). *Character*. Oxford: Oxford University Press.
- May, L., M. Friedman og A. Clark (1996). "Introduction". *Mind and Morals: Essays on Cognitive Science and Ethics*. L. May, M. Friedman og A. Clark (red.). Cambridge, Mass: MIT Press. 1-15.
- Miller, A. (2003). *An Introduction to Contemporary Metaethics*. Cambridge: Polity Press.
- Nagel, T. (1980). "Ethics as an Autonomous Theoretical Subject". *Morality as a Biological Phenomenon: The Pre-suppositions of Sociobiological Research*. G. S. Stent (red.). Berkeley: University of California Press. 198-205.
- Nisbett, R. E. (2003). *The Geography of Thought: How Asians and Westerners Think Differently ... and Why*. New York: Free Press.

-
- Nussbaum, M. C. (2010). "Skills for life". *Times Literary Supplement*. 30.05.2010.
Tilgjengelig på
<http://www.courrierinternational.com/files/illustrations/article/2010/06/1025/NussbaumSkills-Time.pdf> (oppsøkt 25.05.2011)
- Prinz, J. (2008). "Empirical Philosophy and Experimental Philosophy". *Experimental Philosophy*. J. Knobe og S. Nichols (red.). Oxford: Oxford University Press. 189-208.
- Prinz, J. (2009). "The Normativity Challenge: Cultural Psychology Provides the Real Threat to Virtue Ethics". *The Journal of Ethics*. 13: 117-144.
- Railton, P. (1986). "Moral Realism". *The Philosophical Review*. 95:163-207.
- Railton, P. (1995). "Made in the Shade: Moral Compatibilism and the Aims of Moral Theory". *Canadian Journal of Philosophy Supplementary Volume*. 21: 79-106.
- Railton, P. (2004). "Toward an Ethics that Inhabits the World". *The Future for Philosophy*. B. Leiter (red.). Oxford: Clarendon Press.
- Railton, P. (2009). "Two Cheers for Virtue". Notat til seminar ved CSMN. Tilgjengelig på
http://www.csmn.uio.no/events/2009/seminar_w_railton.xml (oppsøkt 06.03.11).
- Ross, L. og R. E. Nisbett. (1991). *The Person and the Situation: Perspectives of Social Psychology*. New York: McGraw-Hill.
- Smith, M. (1994). *The Moral Problem*. Oxford: Blackwell.
- Solomon, R. C. (2005). "What's Character Got to Do with It?". *Philosophy and Phenomenological Research*. 71: 648-655.
- Sorell, T. (1991). *Scientism: Philosophy and the infatuation with science*. London/New York: Routledge.
- Sreenivasan, G. (2002). "Errors about Errors: Virtue Theory and Trait Attribution". *Mind*. 111: 47-68.
- Williams, B. (1985). *Ethics and the Limits of Philosophy*. London: Fontana Press/Collins